



HAL
open science

Resilience analysis and quantification method for 5G-Radio Access Networks

Soumeya Kaada, Marie Line Alberi Morel, Gerardo Rubino, Sofiene Jelassi

► **To cite this version:**

Soumeya Kaada, Marie Line Alberi Morel, Gerardo Rubino, Sofiene Jelassi. Resilience analysis and quantification method for 5G-Radio Access Networks. NoF 2022 - 13th International Conference on Network of the Future, Oct 2022, Ghent, Belgium. pp.1-14, 10.1109/NoF55974.2022.9942669 . hal-03898621

HAL Id: hal-03898621

<https://inria.hal.science/hal-03898621>

Submitted on 14 Dec 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Resilience analysis and quantification method for 5G-Radio Access Networks

Soumeya Kaada, Marie Line Alberi Morel,
Nokia Bell Labs, Nozay, France.

Soumeya.Kaada@nokia.com,
Marie_Line.Alberi-Morel@nokia-bell-labs.com

Gerardo Rubino, Sofiene Jelassi
INRIA/IRISA, Campus de Beaulieu, Rennes, France

Gerardo.Rubino@inria.fr,
Sofiene.Jelassi@irisa.fr

Abstract

A 5G Radio Access Network (5G-RAN) can be disturbed or shutdown due to a variety of failures, in spite of advanced optimization techniques and self-healing methods. Recently, operators started to take an interest in improving the resilience of communication networks with adaptive compensation techniques to mitigate outage-induced performance degradation. However, for the sake of effective and efficient resilience management, it is vital to be able to measure current and prospective resiliency levels of a given 5G-RAN using relevant and explicit metrics. Thus, the characterisation of resilience goes over a thorough analysis of 5G-RAN performance indicators followed by a rigorous quantification of current and future levels of resilience. In this work, we perform an analysis and a quantification of 5G-RAN resilience using a coverage indicator. It is known as a main performance indicator for network planners and operators, coverage is a necessary prerequisite to ensure a certain level of Quality of Service. For that, we model the network coverage using Continuous Time Markov Chains (CTMCs) where coverage status is characterized with multiple states defined with Reference Signal Received Power (RSRP) signal. The proposed Markov model is analytically studied allowing to perform quantitative analysis, predict coverage outage and provide resilience quantification. Using our model, we conduct numerical analysis of several usage scenarios and propose a resilience framework to show the usability of our proposed approach.

Index Terms

Resilience, Coverage, CTMCs, Quantification, Outage prediction, 5G RAN, RSRP

I. INTRODUCTION

5G and beyond networks are expected to provide higher quality of services and to ensure more reliable communications especially in the context of Ultra-Reliable Low-Latency Communication (URLLC) systems or Ultra Dense Networks (UDNs). Given the complexity of the radio environment, multiple intelligent management techniques are deployed in 5G-RAN to ensure the radio performance requirements, like Physical layer Radio Resource Management (RRM) schemes, interference and mobility management algorithms, Coordinated Multipoint (COMP), etc, as well as multiple 5G enabling technologies like Massive Multiple-Input Multiple-output (MIMO) [1]. Moreover, Self-Organizing Networks (SON) concept has been considered to offer 4G and 5G-RAN automatic processes of network management (optimization, configuration, diagnosing and healing). In turn, these Self-Management techniques help engineers to reduce operationally expensive manual oversight and to repair anomalies, degradation or failures like e.g. Coverage and Capacity Optimization (CCO) function. These management mechanisms play a pivotal role to provide an adequate Quality of Service/Experience (QoS, QoE) for end users. However, multiple risk factors can still disturb the radio environment (or RAN) resulting in failure occurrences and poor QoS such as insufficient capacity at BS, fading effects, large variation in antenna performance, energy consumption overload, etc. When the outage occurs, the main ways for engineers to deal with service blackouts are either performing manual repairs or activating self-healing methods [2], which are known to be time consuming. In fact, according to Regulatory Authority for Electronic Communications, Posts and Press Distribution (ARCEP), an average repair time of 46 hours was recorded in 2021 for 85% of quickly detected anomalies [3]. Thus, current efforts in building reliable and optimised Base Stations (BSs), have to go further to protect RAN against disruptive events and service interruption and enhance the network resilience.

Generally speaking, *Network Resiliency* is thought of as the ability of a network to absorb the adverse impacts, adapt to and quickly bounce back from disruptions to continue keeping an acceptable level of service from users' perspective. Traditional resilience solutions used by operators, consisted on deploying costly redundant solutions to address temporally cell outage, like replacing under-performing network equipment with redundant ones. Recently, researchers showed a major interest in developing adaptive mechanisms that enhance resilience. Based on the virtualized infrastructures and the dynamic reconfiguration, prior works proposed to activate Cell Outage Compensation (COC) techniques, as a function of SON, to mitigate outage induced performance in case of failure detection [4]. However, these few developed mechanisms are limited when the outage is hard or impossible to detect. Furthermore, network failures face a lack of management and of anticipation strategy [5]. Besides, most prior works in RAN resilience enhancement propose resilient mechanisms without quantifying the improvements observed with resilience metrics [5]. This makes the majority of resilience solutions in the RAN domain purely qualitative. Thus one of the major challenges addressed in this paper is to provide an analysis and quantification method of resilience in the radio domain to help current or future RAN in risk management. We assume that in the future, service or network providers or operators will need to define the network resilience requirements in Service Level Agreement (SLA) contracts based on target attributes. This implies a clear vision of 5G network resilience and how to manage it quantitatively [5]. The analysis is proposed in the perspective to be part of a resilient management system which its main goal is to supervise the infrastructure and act according to observations. This concept is called proactive resilience [6]. This approach is based on collecting and performing real-time data analytics of system abnormal events to anticipate network blackouts before specifying how to handle and plan for, or how to reduce effects of future disturbances.

Analysing resilience in RAN is related to the analysis of quality of services over the network. The quality of services is characterized by multiple Key Performance Indicators (KPIs) like: good coverage and capacity, low latency, high throughput, availability, etc [7]. Therefore, to obtain an accurate measurement of global service resilience, it is important to integrate all main KPIs in the definition of resilience. To limit complexity and cost of resilience investigation, we adopt a sub-optimal approach that consists of decomposing the initial complex problem into multiple simpler and easier sub-problems. Thus, as a

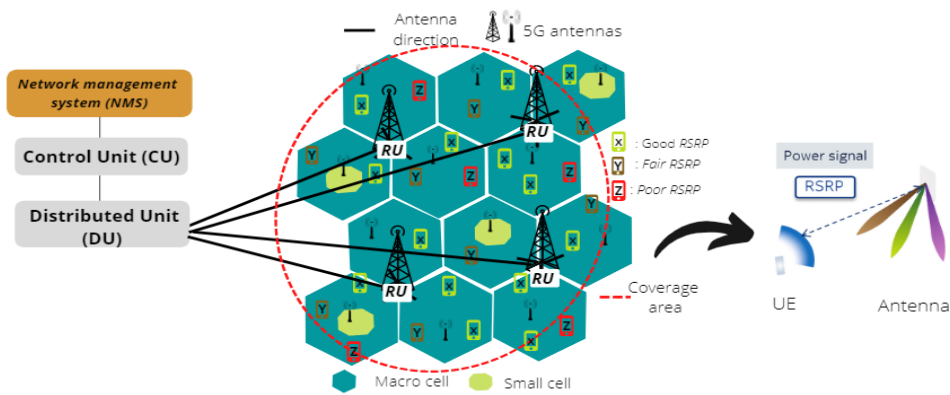


Fig. 1: Radio Coverage scheme and architecture of 5G cellular site

primary approach, we analyze resilience under the scope of coverage metrics that is the main network performance indicator. Coverage is one of the key factors that an operator considers when commercializing cellular communication networks due to its direct impact on service [8]. Besides, coverage failures are hard to troubleshoot and diagnose in practice since logs cannot be retrieved when there is absolutely no coverage [7]. Therefore, an anticipation of coverage outage coupled with quantitative resilience metrics is required.

In this work, we provide a quantitative resilience analysis for 5G-RAN to predict coverage outage and provide resilience quantification. To do this, we are interested in Markov processes to model network coverage states. The multiple states are defined with Reference Signal Received Power (RSRP) signal of the Channel State Information-Reference Signal (CSI-RS). RSRP is a 3GPP standard radio signal [9] and is equivalent to 4G-LTE RSRP. It is a good indicator of the signal strength received by the user and thus of the coverage status. We define and discuss several candidate metrics of resilience using the coverage performance indicator. A numerical analysis of several types of networks is conducted in order to study and discuss the obtained results coming from a variety of RAN use-cases. Finally, an overview of a RAN resilience framework is discussed to leverage the proposed modeling and resilience analysis within a 5G RAN system architecture.

The paper is structured as follows. Section II discusses related work about resilience and proactive analysis in the RAN domain. Section III contains a brief resilience overview in the RAN domain and presents the 5G radio coverage use case. In Section IV we present resilience analysis using Markov processes for modelling network coverage, and we define relevant resilience metrics for our quantitative analysis. In Section V, we discuss numerical results in different scenarios and in Section VI we detail the RAN resilience framework. Conclusions are given in Section VII.

II. RELATED WORK

The first and maybe crucial step of resilience management is known as resilience analysis [5]. The analysis allows engineers to characterize the system behaviour via modeling techniques, and to quantify the resilience using suitable metrics [10]. In the literature, we noticed that most works seeking to enhance resilience propose reactive mechanisms without quantifying the improvements. We refer the reader to works on compensation methods of SON for coverage issues [4, 11]. For example, [4] proposes a COC method that mitigates outage impacts and restore coverage via controlling different radio parameters such as antenna tilt and the up-link target received power level P_0 . However, in the absence of an analysis and anticipation of disruptive events, these mechanisms are limited when facing failures that are difficult to detect, such as sleeping cells [2], and remain purely qualitative without resilience quantification.

To analyse resilience, it is interesting to get inspired from dependability studies¹ that are well elaborated

¹Dependability and reliability methods are designed to avoid failures and to be mainly active before outage occurrence (e.g predictive maintenance).

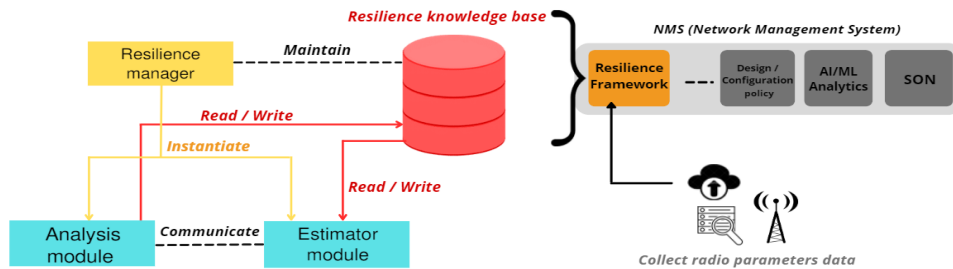


Fig. 2: RAN Resilience Framework

for the analysis of repairable systems than resilience discipline². For instance, a previous work on reliability analysis was proposed in [12]. The authors used Continuous Time Markov Chains (CTMCs) to model the status of a specific BS into 3 states (optimal, sub-optimal and outage), and extract reliability metrics from the model. Through a fault prediction framework, they analyse the transient and steady state probability distributions of 3 study cases to predict network status. Based on metric values the framework enhance reliability via activation of root cause analysis to avoid outage. However, the reported model captures only high-level features of many categories of failure events affecting a BS, and neglects the complex cellular network environments. Besides, only reliability metrics are computed from the model. It has to be noted, that quantifying system resilience is more complex than what implies simple dependability attributes like reliability known metrics such as Mean-time-to-failure (MTTF) [13]. It is not a simple Boolean function or a binary view of the system, it is always a matter of degree, and for sure system resilience is typically not measurable on a single ordinal scale [13].

In RAN, there is no common and well-defined resilience concept which explains why we find many variations of the idea adapting it to each case [14, 15]. In [16], the authors proposed a definition of availability, resilience and reliability in the RAN domain. Resilience has been expressed as an integral of availability of hardware and radio resources for service, over a period of time. Others propose to compute survivability³ (a sub-field of resilience) in wired networks, in terms of the loss probability or the delay distribution of non-lost packets in wired networks [18].

Unlike previous work, we propose modelling using CTMCs for coverage states instead of considering at once all failures on BS [12]. We consider radio data when modelling the network coverage, to reflect as much as possible the reality of the radio environment in the proactive analysis. We then adapt existing measures and define new ones to quantify resilience in terms of coverage. We think of resilience metrics as an attribute uncorrelated with other reliability or survivability metrics. To the best of our knowledge, this is the first work that addresses resilience quantification using coverage indicators extracted from CTMCs in the RAN domain. Besides, our analysis method can be useful to combine with reactive mechanisms like COC methods using our framework.

III. RAN RESILIENCE CONTEXT

The management of proactive resilience in RAN may be separated into: a first part that defines methods for proactive system analysis, and a second part that specifies the resilient mechanisms to fulfill the requirements in QoS. In this section, we give an overview of the general resilience management scheme in a 5G architecture and focus on the resilience analysis part for coverage issues. In section VI a detailed framework of the resilient management scheme is discussed.

A. 5G cellular network architecture

Consider a 5G network comprising multiple macro cells (MCs) and a large number of small BSs. Such network deploys an architecture with large small cells and handles multiple heterogeneous devices. UEs are connected to Remote Radio Heads (RRH) or Radio Units (RU), which sends information to 5G

²Resilience methods try to adapt the network once the outage occurs. They mainly react after failure to quickly restore the service.

³Survivability focus on attacks and large scale disasters [17]

network via the Distributed Unit (DU) and then to the Control Unit (CU) (see figure 1). The management plan of the radio system is executed by the Network Management System (NMS). This module can integrate a resilience management framework. Depicted in figure 2, the framework gives an overview of how proactive resilience can be supported in 5G-RAN or in future cellular network architecture. It can be applied on demand to enhance resilience of a group of cells, a geographical area or a slice. It comprises two principle modules. A first module responsible for analysis of resilience. Using network performance indicators the analysis module allows predicting outage and quantifying the impacts. The second estimator module uses the results of the resilience analysis to perform decision making on what type of adaptive resilience mechanisms to activate. The framework works using data collected from BSs and stores all the relevant information needed for the process into a knowledge base. The framework will be further discussed in section VI to demonstrate its usability to enhance resilience of coverage aspect. The purpose of this work is to provide a method for the analysis module while the estimator module is left as a future work.

B. 5G cellular network use case: radio coverage

In this section, we discuss about radio coverage of a given group of serving cells deployed in a given geographical area. The area is considered as covered when the User Equipment (UE) is able to connect to their mobile network, establish and maintain a call for a minimum period of time, achieve a specific data transmission speed, and access different services. An UE can connect to network if it receives a signal strength from BS that is above a minimum acceptable coverage threshold. In urban areas, UE can be covered simultaneously from multiple gNodeBs to reduce coverage related-failures. But, rural areas have sparse deployments where no or rare overlapping coverage is available so coverage problems are more frequent. In both cases, coverage degradation or loss can still be observed within a geographical area. For 5G, the CSI - radio received signal $RSRP$ is a valuable layer 1 measurement to provide information about the signal strength received by users within a service area. It is defined as a linear average over the power contributions (Watts) of a single Reference Signal (RS) resource element (in dBm) [9]. The measures lie between -140 dBm and -44 dBm and represent a good metric of the radio coverage experimented by users.

In our work, we assume a coverage area in which a set of phone users are distributed uniformly. The number of users within the area is considered static. In real applications it may correspond to a constant traffic during a fixed period of time where the number of users vary slowly. Under the hypothesis of static traffic and uniform distribution of users, we split users into 3 categories according to their levels of $RSRP$. Figure 1 shows in the 5G cellular architecture the positions of each category users in terms of $RSRP$ level that can be observed over a coverage area.

Let us denote by $x = x(t)$, $y = y(t)$ and $z = z(t)$ the percentage of users that has *Good*, *Fair*, *Poor* levels of $RSRP$ respectively at t . The traffic is considered static and the sum of all percentages is equal to 100%. We investigate 2 configurations of $RSRP$ levels (C1 & C2). Table I shows the radio RF conditions of $RSRP$ values in both cases 1 & 2. The first type corresponds to a network with minimal requirements in terms of coverage, whereas the second one corresponds to networks with higher requirements in terms of coverage, and thus higher $RSRP$ values boundaries.

	Classification	(1) $RSRP$ (dBm)	(2) $RSRP$ (dBm)
RF Conditions	Good (or excellent)	-90 to -44	-80 to -44
	Fair (Mid cell)	-126 to -91	-100 to -81
	Poor(Cell Edge)	-140 to -127	-140 to -101

TABLE I: Radio RF condition configurations ($RSRP$)

IV. COVERAGE BASED RESILIENCE ANALYSIS

In this section, we present the analysis of the radio system resilience using CTMCs, a powerful tool widely used in repairable systems study. It is based on stochastic processes allowing to predict network

status and characterizing system behaviour.

Methodology for resilience analysis and quantification

- (A) Definition of a Markov model (states and transitions) to characterize the behaviour of the coverage status within a service area.
- (B) Prediction of outage occurrence while computing transient and steady state probability distributions.
- (C) Quantification and analysis of system resilience using resilient metrics based on coverage performance indicator.

A. Model Development

We propose a CTMC model to represent the different coverage states of L cells of 5G in a given geographical area. The model will serve to analyse the network resilience. It is formed by states and transitions between these states. The CTMC model relies on $RSRP$ measurements that are collected within RAN to track network coverage evolution.

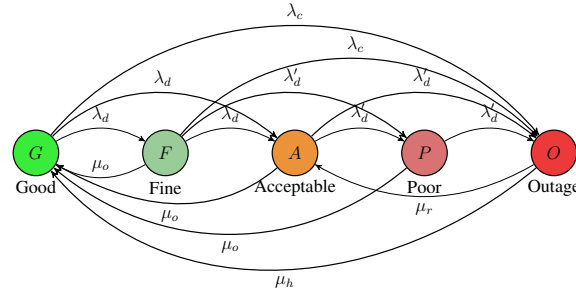


Fig. 3: Markov model with resilience solution

Let $X(t)$ be a CTMC with 5 states (see Figure 3) with state space $S = \{G = \text{“Good”}, F = \text{“Fine”}, A = \text{“Acceptable”}, P = \text{“Poor”}, O = \text{“Outage”}\}$. These states are defined to reflect the behaviour of the radio coverage observed by engineers with a fine grained degrees. Each state is defined in terms of the tuple (x, y, z) previously described within the coverage area of a gNodeB. The model considers the radio system to be resilient when the majority of users within a service area satisfies a minimum acceptable coverage threshold specified with $RSRP$ levels given in Table I. We set out to 5% the admissible percentage of users, which receive $RSRP$ below the threshold associated with poor $RSRP$. This implies that the operator network guarantees a minimal radio coverage of 95% to be resilient. It is directly inspired from assumptions of mobile network radio planning for which the term of location probability is introduced. It describes the probability of the receiver being able to capture the signal with a power higher than the receiver sensitivity. On the contrary, a location probability in outdoor area of below to 95% is qualified as a non admissible coverage [19].

The coverage model of L cells is made of 5 states (3 main states: G, A, O and 2 intermediate states: F, P) as follows:

- 1) $X(t) = G$ for “Good” - Cells carrying traffic with best performances. A coverage area is Good and only if the percentage of users with poor $RSRP$ is less than 5% and the percentage of users with good $RSRP$ is the highest one, i.e, iff $x > y, x \geq z, x > y + z$ and $z < 5\%$.
- 2) $X(t) = F$ for “Fine” - Cells carrying traffic with fine performances. This intermediate state is quite similar to G state where percentage of users with good $RSRP$ is the highest one i.e, iff $x > y, x \geq z$ and $z < 5\%$ but $x \leq y + z$.
- 3) $X(t) = A$ for “Acceptable” - Cells still carrying some traffic, but with slightly lower performances. A cell coverage is acceptable if and only if the percentage of users with poor $RSRP$ is between $0 \leq z < 4\%$ and the percentage of users with fair $RSRP$ is the highest one, i.e $y > x, y \geq z$
- 4) $X(t) = P$ for “Poor” - Cells carrying on some traffic, but with poor performances. This intermediate state is observed when percentage of users with poor $RSRP$ is between $4 \leq z < 5\%$ and the percentage of users with fair $RSRP$ is the highest one, i.e $y > x, y \geq z$.

5) $X(t) = O$ for “*Outage*” - A severely degraded coverage or with no traffic carried. A coverage outage is observed if and only if the percentage of users having poor *RSRP* is greater than 5%, i.e, $z \geq 5\%$.

The classification of coverage status in 5 states allows having a fine grained observation and participates to a better anticipation of outages. The reader can verify that all states are fulfilled by the tuple (x, y, z) , where the sum of all percentages is 100%, and the coverage status of the service area can only have one state at a time without correlation. In the proposed Markov model, we classify failures as follows:

- *Degradation*: perturbations (e.g. interference or inaccurate antenna downtilt) that degrade good or fine coverage states to sub-optimal states A or F without causing failures. We denote the rate of their occurrence by λ_d .
- *Severe degradation*: perturbations (e.g. sleeping cell due to failures) that cause likely the cell outage, and are severe than simple degradation in terms of the resulting states (P or O). It may occur at sub-optimal states F,A,P. The corresponding transition rate is λ'_d .
- *Critical failures*: major and unpredictable cell outages occurring less frequently. It might correspond to disasters, due to environmental conditions (hurricanes, ..) or human made (cable/power cut, ...) etc. The associated transition rate is λ_c .

We note that similar transitions rates appear between several pairs of states like λ_d , because the same types of disruptive events are assumed to be more likely to happen. We categorize recovery transitions in the model into 3 types:

- *Optimization mechanisms*: are the currently applied coverage enhancements techniques in RAN like Access Traffic Steering, switching and Splitting (ATSSS) used in overlapping coverage with carrier aggregation or any other technique proposed in Rel-15/17 of 5G. They are responsible for increasing coverage quality from state (F, A, P) to G due to degradation. These techniques are practically not designed to face major and severe failures. The corresponding transition rate is μ_o . We assume that the same types of mechanisms are applied in F, A, P, so in average, it takes the same speed μ_o to switch to the state G. The absence of μ_o between intermediate states is explained with the fact that optimizations are theoretically responsible to restore network status to an optimal and not sub-optimal or intermediate states.
- *Healing mechanisms*: restore coverage from state O to G through total healing/repairs or self-healing mechanisms. It may also involve adaptive resilient mechanisms that mitigate failures in an optimal way to not restrict resilience in its basic definition. The transition rate is μ_h .
- *Resilience mechanisms*: restore coverage from state O to A. In fact, the main and by default goal of resilience is to maintain an acceptable level of service under failures. These mechanisms are essentially implemented by adaptive resilient mechanisms (i.e. alternative systems, redundancy, rollback/roll-forward functions, compensation or mitigation functions, failure isolation, ...). We denote the associated transition rate by μ_r .

The infinitesimal generator matrix \mathbf{Q} of $X(t)$ is:

$$\mathbf{Q} = \begin{bmatrix} -[2\lambda_d + \lambda_c] & \lambda_d & \lambda_d & 0 & \lambda_c \\ \mu_o & -[\mu_o + \lambda'_d + \lambda_d + \lambda_c] & \lambda_d & \lambda'_d & \lambda_c \\ \mu_o & 0 & -[2\lambda'_d + \mu_o] & \lambda'_d & \lambda'_d \\ \mu_o & 0 & 0 & -[\lambda'_d + \mu_o] & \lambda'_d \\ \mu_h & 0 & \mu_r & 0 & -[\mu_h + \mu_r] \end{bmatrix} \quad (1)$$

To investigate the case of an infrastructure with no resilience mechanisms, the same model can be used except for μ_r which is set to $\mu_r = 0$. In the following, the transient and the steady state distribution equations are deduced from the model in order to compute state probabilities, to predict the network coverage outage and status and to define resilience metrics.

B. Transient and steady state analysis

Transient state analysis: Using the generator matrix \mathbf{Q} , the dynamic behaviour of the resilient CTMC model can be described by the Kolmogorov differential equation $P'(t) = P(t)Q$ where $P(t) = (P_G(t), P_F(t), P_A(t), P_I(t), P_O(t))$.

and $\sum_{i \in S} P_i(t) = 1$. The transient state probability vector $P(t)$ of the resilient model can be obtained by solving the differential equations system defined as follows:

$$\begin{cases} P'_G(t) = -(2\lambda_d + \lambda_c)P_G(t) + \mu_o P_F(t) + \mu_o P_A(t) \\ \quad + \mu_o P_P(t) + \mu_h P_O(t), \\ P'_F(t) = \lambda_d P_G(t) - (\lambda_d + \lambda'_d + \lambda_c + \mu_o)P_F(t), \\ P'_A(t) = \lambda_d P_G(t) + \lambda_d P_F(t) - (2\lambda'_d + \mu_o)P_A(t) \\ \quad + \mu_r P_O(t), \\ P'_P(t) = \lambda'_d P_F(t) + \lambda'_d P_A(t) - (\lambda'_d + \mu_o)P_P(t), \\ P'_O(t) = \lambda_c P_G(t) + \lambda_c P_F(t) + \lambda'_d P_A(t) \\ \quad + \lambda'_d P_P(t) - (\mu_h + \mu_r)P_O(t). \end{cases} \quad (2)$$

Steady state analysis: The steady state distribution reflects the behaviour of the system in the long run. It requires to solve the linear equation system $\pi Q = 0$ where $\pi = (\pi_G, \pi_F, \pi_A, \pi_P, \pi_O)$, plus the normalization condition $\sum_{i \in S} \pi_i = 1$. The linear system is:

$$\begin{cases} -(2\lambda_d + \lambda_c)\pi_G + \mu_o\pi_F + \mu_o\pi_A + \mu_o\pi_P + \mu_h\pi_O = 0, \\ \lambda_d\pi_G - (\lambda_d + \lambda'_d + \lambda_c + \mu_o)\pi_F = 0, \\ \lambda_d\pi_G + \lambda_d\pi_F - (2\lambda'_d + \mu_o)\pi_A + \mu_r\pi_O = 0, \\ \lambda'_d\pi_F + \lambda'_d\pi_A - (\lambda'_d + \mu_o)\pi_P = 0 \\ \lambda_c\pi_G + \lambda_c\pi_F + \lambda'_d\pi_A + \lambda'_d\pi_P - (\mu_h + \mu_r)\pi_O = 0. \end{cases} \quad (3)$$

To obtain the transient and steady state probability vectors of a network that doesn't integrate resilient mechanisms, the differential equations system (2) and linear equations system (3) should be solved after setting $\mu_r = 0$.

C. Resilience metrics using coverage performance indicator

The problem of quantifying system resilience has been elusive [10]. There is a lack of standardization in resilience quantification especially in the RAN domain, and many researchers are coming up with their own definitions of resilience metrics. For a proactive resilience analysis, we investigate metrics that enable an estimation of future changes in the network (see part (1)). In fact, we propose two stochastic metrics that quantify the resilience level. They allow to measure the distance between the estimated coverage status and the targeted minimal coverage requirements. In part (2), we propose a metric that assesses, how the network can ‘‘bounce back’’ to a normal state via various Resilience Mechanisms (RMs). We refer it as a score of resilience. So, It is a measurement of the improvement in term of resilience brought by RMs.

(1) Estimating the evolution of resilience in terms of coverage within a geographical zone over time dependable metrics:

– We define resilience metric, referred as $\mathcal{M}(t)$, as the global stochastic mean value of RSRP medians at time t :

$$\begin{aligned} \mathcal{M}(t) = & med_{Good_{RSRP}}(t)P_G(t) + med_{(Good+Fair)_{RSRP}}(t)P_F(t) + med_{Fair_{RSRP}}(t)P_A(t) \\ & + med_{(Fair+Poor)_{RSRP}}(t)P_P(t) + med_{Poor_{RSRP}}(t)P_O(t). \end{aligned} \quad (4)$$

This metric considers the probability of being in each state $P_i(t)$ with $i \in \{G, F, A, P, O\}$, which is obtained from the analysis of the Markov chain. Also, $med_{\theta_{RSRP}}(t)$ where $\theta \in \{Good, Fair, Poor\}$ represent the RSRP median of state Good(x), Fair(y) and Poor(z) described in Table I. According to state definition, we consider only the population of users having in majority good or fair RSRP for computing the RSRP median in G, F and A states. Instead, we consider the population of users having in majority fair or poor RSRP for computing the RSRP median in states P and O states

– Based on [16], we propose another resilience metric that measures the availability of an acceptable coverage over a period T_c . It's written as an integral over a duration T_c around time t of RSRP above a minimal acceptable threshold denoted as γ as follows:

$$\mathcal{P}R_{T_c}(t) = \frac{1}{T_c} \int_{t-T_c/2}^{t+T_c/2} \sum_{S_R} \mathbb{P}(RSRP > \gamma) dt, \quad (5)$$

T_c is a sliding period considered to track coverage availability, and S_R is the set of states $\{G, F, A, P\}$ where most $RSRP$ values are greater than a given threshold γ to have fair or good $RSRP$ level.

(2) Quantifying the gain that might be brought by a network integrating a RM:

- We denote the resilience value by $RM_Res(t)$ when the network integrates RM and by $NRM_Res(t)$ when there is no RM. We define the metric of resilient score as the ratio between the resilience level between RM and NRM cases. The score inspired from a previous work in [10] in equation (6). It is expressed as follows:

$$Res_score(t) = \frac{RM_Res(t)}{NRM_Res(t)} \quad (6)$$

This metric is general but in this paper, $RM_Res(t)$ and $NRM_Res(t)$ will be computed using equations (4) or (5) and expressed in dB.

Remark about methodology and resilience metrics

1) Once the transition rates in the model are available, the analyst can answer “what if” questions and evaluate any kind of metric, even those that will be defined “tomorrow”. For instance, consider, as a performance criteria, the time spent at the set of main states $\{G, A\}$. Call τ that random variable, assuming the model starts at G . Standard Markov properties give us the queue of the distribution of τ (a kind of reliability at t metric), or its average. As an illustration, the expectation $\text{Exp}(\tau)$ is:

$$\text{Exp}(\tau) = \frac{\lambda_d + 2\lambda'_d + \mu_o}{(2\lambda_d + \lambda_c)(2\lambda'_d + \mu_o) - \lambda_d\mu_o}.$$

The same can be done if we look at the subset of states $\{G, F, A\}$, or if we are interested in other more complex metrics. We don't pursuit this line here, for lack of room.

2) For sure, other resilience metrics can also be defined to enrich the analysis. Our work fits a first step of the whole resilience vision. We are currently in the steps of investigating resilience metrics considering only a coverage performance indicator. In future work, another category of metrics enriched with multiple QoS indicators as data transmission speed or latency will be considered. Their inclusion is more appropriate to ensure an accurate resilience measurement and to enhance risk management during decision making of RM solutions.

V. NUMERICAL RESULTS AND DISCUSSIONS

For numerical analysis, we apply our methodology on four practical study cases that depict different configurations of transient rates. Firstly, we discuss about the proposed rate values for each study case given in Table II. Then, the rate values are used to analyse transient and steady state probabilities given in equations (2) and (3). They help to predict outage occurrence and quantify the resilience. They are drawn in figures 4 and 5 in the case of basic and critical networks, respectively. Then, the evolution of resilience over time, given in equations (4) and (5), is studied quantitatively. It is shown in figures 6 and 7(a), respectively. In figure 7(b), we analyse the score of resilience (eq. 6) that quantifies the improvements brought by RMs in figure 7.

(a) **Study cases configurations:** We consider 4 study cases that correspond to various real network types. We distinguish 2 types of network groups that are *basic network* and *critical network*. The first group corresponding to cases 1 and 2 represent common networks that carry basic network services with minimal levels of QoS requirements. The second group corresponding to cases 3 and 4 represent

distinguished networks that carry critical services with higher requirements in terms of QoS. Critical networks are currently well optimized against degradation and apply frequently preventive dependability methods to reduce failure events. To evaluate the impact of resilience, we assume that each type has two different resilience configurations. Cases 1 and 3 don't integrate resilience schemes while cases 2 and 4 have RM enabled. Table II contains theoretical values in hours of the failure transient rates λ_i and recovery values μ_i for all study cases. They are set based on analysis of real RAN performance problems reported in works [3, 12]. In real implementation, these rates are computed statistically from failure log data.

Transient rates/ Configuration cases	Basic Network Services		Critical Network services	
	Case 1 No Res	Case 2 With Res	Case 3 No Res	Case 4 With Res
λ_c ($hour^{-1}$)	1/1440	1/1440	1/2140	1/2140
λ_d ($hour^{-1}$)	1/10	1/20	1/20	1/20
λ'_d ($hour^{-1}$)	1/24	1/24	1/72	1/72
μ_o ($hour^{-1}$)	1/5	1/4	1/4	1/4
μ_h ($hour^{-1}$)	1/48	1/24	1/48	1/24
μ_r ($hour^{-1}$)	0	1/2	0	1

TABLE II: Transient rates of all study cases

(b) **Discussion on transient values:** Table II shows that the rate λ_c of major failures is observed in average once every 2 months for basic network services, and around every 3 months for critical network services. In case 1, where no resilience is enabled, we set the degradation rate λ_d to 1 every 10 hours and severe degradation λ'_d , that can likely cause the outage, every 24h. To limit degradation, optimization functions operate in background. We assume a recovery transition happens with rate 1/4h, in other words, those optimization functions take, on the average, 4h to restore coverage (and $\mu_o = 1/5h$ for case 1). In case of outage occurrence, healing and repairs take roughly 2 days including the detection time and the root cause analysis either automatically or manually (that is, $\mu_h = 1/48h$). In case 2 where the network uses RMs, we should expect observing less failure occurrences and faster recovery processes than in a non resilient network. In fact, we assume it takes 2h on average to deploy an adaptive resilience solution to restore the coverage to an acceptable level ($\mu_r = 1/2h$). This would increase occupancy time at state A where optimisation is applied efficiently within 4h. It has to be noted that RM activation impacts the network through decreasing slightly rates of degradation or severe degradation. Here, we choose to consider it by decreasing λ_d to 1/20h to illustrate the impact of RMs on degradation transitions. Rate μ_h is also increased compared to case 1 because it integrates total repairs plus optimal resilience solutions as explained in Section IV. For the second type of networks, we assume that we will observe less failure events in critical infrastructures compared with the first type since they already include rigorous optimizations and more efficient healing functions, but still not sufficient to cope with all failures encountered, and sometimes, it is possible that they don't have adaptive resilient solutions to face quickly the failures (case 3). Therefore, we increase λ'_d to once every 3 days, the arrival time of degradation λ_d to 20h and we assume that the optimizations take in average 4h. These values are identical for both cases 3 and 4. However, in case 4 we focus on studying the impact of RM activation on the resilient transition μ_r only set to 1h and the optimal resilient solution or total healing transition μ_h that takes in average 24h.

(c) **Outage prediction over a time period:** Figures 4 and 5 represent the results of transient analysis of states probabilities $P(t)$. The probabilities are computed over 15 days in all cases. The analysis of probabilities reveals that a basic network with no resilience (case 1) encounters a higher failure risk measured by $P_O(t)$ than in the other cases with approximately 40% after 5 days, while for critical networks with no resilience (case 3) its value is around 10%, due to the optimisations assumed to occur in this type of network. In addition, introducing resilience reduces P_O to 2-5% in case 2 and to 1% in case 4, and enhances performances through increasing P_G . We also notice that P_A is more stable after 2 days in case 2 and doesn't decrease compared with case 1, meaning that resilience creates more chances to be in state A and increases the occupancy time as highlighted above. CTMCs are powerful tools to track short term network evolution. For instance, the evolution of $P_O(t)$ allows the operator to predict the

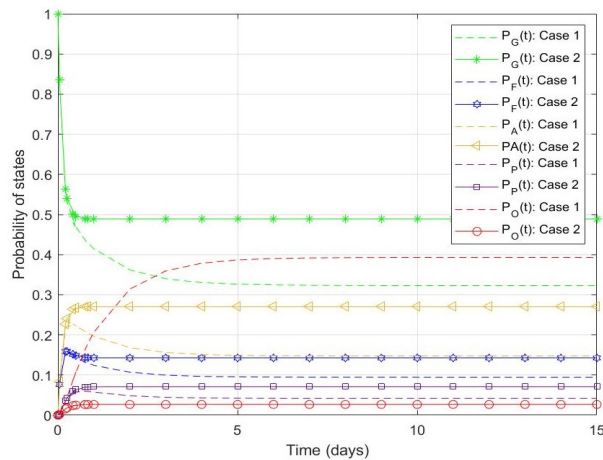


Fig. 4: Transient analysis of cases 1 and 2

outage risk occurrence within a time limit, e.g 5 days. This is helpful to anticipate correctly the disruption according to operator's resilience policy. For example, resilience solution like adaptive RM (e.g. antenna tilt modification) can be prepared for deployment and at a suitable activation time.

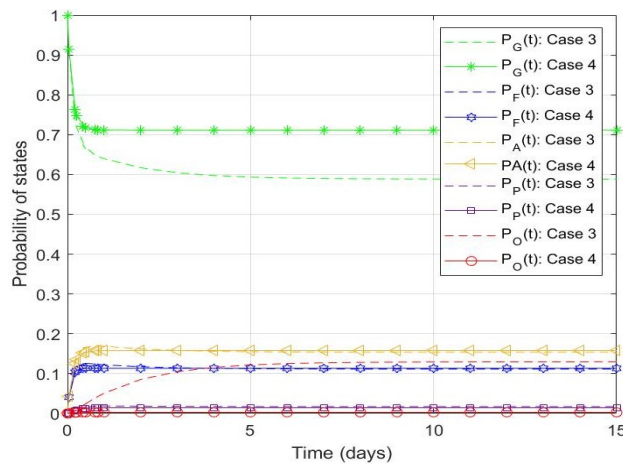


Fig. 5: Transient analysis of cases 3 and 4

(d) **Quantitative evolution of resilience:** The metrics defined in section IV-C quantify resilience level and gain to help operators or network owners getting useful and familiar resilience measurements insights. We assume that RSRP are uniformly distributed for both RSRP configurations of table I and we plot

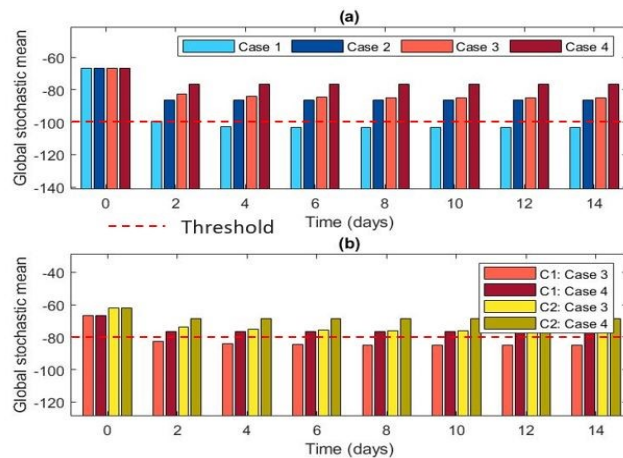


Fig. 6: Global stochastic mean of medians of $RSRP$ (dBm) for: (a) First $RSRP$ configuration for all cases (C1)- (b) Both $RSRP$ configurations (C1 and C2) for critical networks

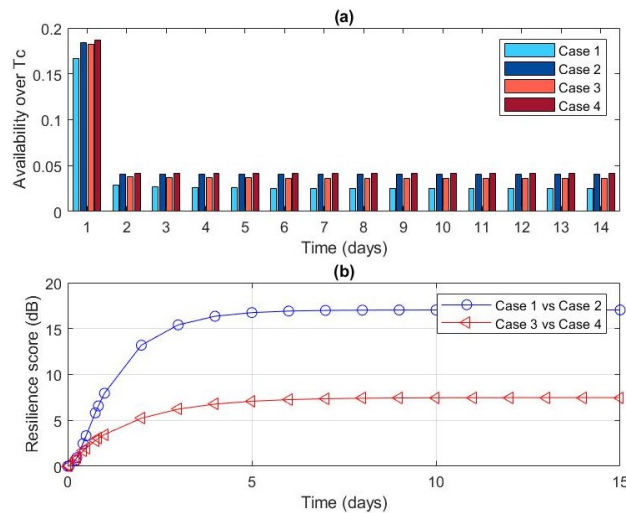


Fig. 7: **(a)** Acceptable coverage availability over a sliding period $T_c=2d$ **(b)** Basic and critical networks resilience score.



Fig. 8: Proactive resilience analysis and estimation steps

the evolution of global stochastic mean of RSRP medians in figure 6. To observe how this metric behaves we use (a) the same RSRP configuration C1 for all study cases and (b) different configurations C1 & C2 on same study cases (we focus on case 3 and 4). The observation gives relevant information for network planners or operators on how RSRP levels decrease over time under potential failures, and what are minimal resilience conditions to ensure. These conditions may serve to be compared with a threshold (eg. (a):-100 dBm, (b): -80 dBm) predefined within an operator policy, to anticipate and avoid coverage outage. We observe that in the case of basic network with no resilience (case 1) the network resilience is under the threshold fixed by policy once every 4 days. In figure 6(b) we notice different values of resulting global stochastic mean of RSRP medians on same cases of critical networks using different configurations. According to this figure, we verify that the critical network with RM (case 4) is the more resilient. Similarly, we can observe in figure 7(a) the availability over a sliding period of 2 days for all cases. We notice that the availability falls after 48h to around 5%. This metric gives another vision of the resilience evolution in the future than the transient probabilities computed at instant t . Finally, figure 7(b) shows resilience score to quantify the improvements of basic and critical networks with RM. The score is computed using global stochastic mean of RSRP medians. We note an optimal improvement of RSRP with 17dB in basic networks and 7dB in critical networks.

To conclude, proactive resilience analysis can be done through Markov modeling and quantification using various metrics for short term prediction. The usage of average RSRP over the coverage zone allows focusing on the impact of failures instead of their types, and hence all types of failures are considered. The analysis quantifies how much vulnerable the basic networks with non resilient solutions are, and how the RM solutions improve network coverage performance status either in basic network services or strengthen more critical network services. However, CTMC is still limited for the long run and looking on RSRP signal only for providing indication on coverage status is not enough. Thus, other signals and KPIs may be considered (RSRQ, throughput...). Besides, accurate resilience measurement requires to consider other indicators like interference, capacity, etc, to quantify the impact of failures on multiple service aspects. In real application, the challenge is to obtain the correct transient rates from past failure logs of the observed networks. Next section provides an overview of utilisation of this method in RAN.

VI. DATA-BASED RESILIENCE FRAMEWORK

As discussed in Section III, a resilience management scheme can be deployed to enhance RAN resilience within a control loop via a data-based and automated framework located at NMS level in 5G (figure 2). In the case of coverage, figure 8 shows that the framework starts with the analysis module learning the transient rates (λ_d, μ_r, \dots) using ML tools or statistical learning from past failure logs to build the Markov model. These estimated rates will be used to analyse the transient and steady state distribution to predict the network status and quantify resilience. Based on the output of the analysis and the resilience metrics, some scripts containing instructions can be prepared to activate adaptive RMs to restore coverage on the area, such as modifying dynamically antenna tilt or PO of neighboring cells for Outage compensation [4]. The estimator module is driven by the operator resilience policy, to take the right decisions on which RM to activate and at what time. The estimator can use intelligent decision algorithms like regression or reinforcement learning. Both modules run in background and in real time, and can communicate with each other. The historical data, model, transient rates, policies and scripts are stored into a database (Figure 8). Besides, a resilience manager will serve to instantiate the modules, coordinate their communication and control the process. Finally, the resilience process is evaluated with real-time metrics to validate the accuracy of the estimation process. The pipeline is executed within a life cycle.

VII. CONCLUSION

This paper provides a proactive quantitative resilience analysis for 5G-RAN. The analysis uses Markov processes for modelling network coverage using RSRP data. Based on the model, we predict coverage outage occurrence and propose resilience metrics to quantify the level of network resilience. To show the usability of our approach, we present a numerical analysis in 4 different use case scenarios. A data-based resilience framework is proposed. According to numerical results, the proactive analysis can help to characterize the behaviour of the system in short or mid-term, and take anticipated decisions to improve the system by activating resilience mechanisms. These resilience mechanisms envision to improve current basic networks and strengthen the already well optimized critical networks. For future work, the proposed predictive model will be tested on real or simulated radio data to compute the transient rates. It will be also extended to consider other key performance indicators (throughput, latency, etc). Finally, machine learning methods will be considered to help decision making in resilience mechanisms activation.

REFERENCES

- [1] Mary A. Adedoyin and Olabisi E. Falowo. Combination of ultra-dense networks and other 5g enabling technologies: A survey. *IEEE Access*, page VOL. 8, 2019.
- [2] M. Firdaus, S. Haryadi, and W. Shalannanda. Sleeping cell analysis in lte network with self-healing approach. In *IEEE TSSA*. IEEE, 2019.
- [3] Autorité de régulation des communications électroniques et des postes (ARCEP). "arcep annual report", 2021.
- [4] M. Amirijoo, L. Jorgueseski, R. Litjens, L.C. Schmelz. Cell outage compensation in lte networks: Algorithms and performance assessment. In *IEEE Conference on Vehicular Technology*. IEEE, 2011.
- [5] A. Mottahedi, F. Sereshki, M. Ataei, A. N. Qarahasanlou, and A. Barabadi. Resilience analysis: A formulation to model risk factors on complex system resilience. In *The Society for Reliability Engineering, Quality and Operations Management (SREQOM)*, 2021.
- [6] S. Jackson and T. L. J. Ferris. Proactive and Reactive Resilience: A Comparison of Perspectives. In *Insight magazine of the International Council on Systems Engineering (INCOSE)*, 2015.
- [7] I. Stoica A.P Iyer, L.E Li. Automating diagnosis of cellular radio access network problems. In *MobiCom*. MobiCom, 2017.
- [8] Réseau francophone de la régulation des télécommunications (FRATEL). "measuring mobile network performance: Coverage, quality of service and maps.", 2020.
- [9] Evolved universal terrestrial radio access (E-UTRAN); requirements for support of radio resource management, 2018.

- [10] J.E. Ramirez-Marquez D. Henry. Generic metrics and quantitative approaches for system resilience as a function of time. *Reliability engineering and system safety*, pages 114–122, 2012.
- [11] I. de la Bandera, P. Munoz, I. Serrano, R. Barco. Adaptive cell outage compensation in self-organizing networks. *IEEE Transactions on Vehicular technology*, pages vol. 67, n. 6, 2018.
- [12] Ali Imran Hassan Farooq, Md Salik Parwez. Continuous time markov chain based reliability analysis for future cellular networks. In *IEEE Globcom*. IEEE, 2015.
- [13] A. Avizienis, J.C. Laprie, B. Randell, and C. Landwehr. Basic concepts and taxonomy of dependable and secure computing. *IEEE Transactions on Dependable and secure computing*, page Vol 1. NO.1, 2004.
- [14] Craig G Rieger. Resilient control systems: Practical metrics basics for defining mission impact. In *International symposium on Resilient control systems*. IEEE, 2014.
- [15] Daniel J. Rosenkrantz, S. Goel, S. Ravi, J. Gangolly. Structure-based resilience metrics for service-oriented networks. In *5th European Dependable Computing Conference*. EDCC, 2004.
- [16] Ilaria Malanchini Vinay Suryaprakash. Reliability in future radio access networks: from linguistic to quantitative definitions. In *International Symposium on Quality of Service (IWQoS)*. IEEE, 2016.
- [17] J.P.G. Sterbenz, D. Hutchison, and E.K. Çetinkaya. 'Resilience and survivability in communication networks: Strategies, principles, and survey of disciplines'. *The International Journal of Computer and Telecommunications Networking*, 2010.
- [18] Poul E. Heegaard and Kishor S. Trivedi. 'Network survivability modeling'. *The International Journal of Computer and Telecommunications Networking*, 2009.
- [19] Ajay R. Mishra. 'Advanced Cellular Network Planning and Optimisation: 2G/2.5G/3G...Evolution to 4G'. Wiley, 2007.