LTE-ADVANCED PRO

A Unified Traffic Steering Framework for LTE Radio Access Network Coordination

Marcin Dryjanski and Michal Szydelko

The authors present an overview of recent 3GPP developments in the area of traffic steering that are shaping the road toward the approaching 5G systems. Due to the increasing availability of RAN features addressing mobile data delivery over a variety of spectrum access techniques, a new traffic steering design is required for the evolving LTE-Advanced Pro standard.

ABSTRACT

This article presents an overview of recent 3GPP developments in the area of traffic steering that are shaping the road toward the approaching 5G systems. Due to the increasing availability of RAN features addressing mobile data delivery over a variety of spectrum access techniques, a new traffic steering design is required for the evolving LTE-Advanced Pro standard. LTE evolution toward 5G brings an opportunity to introduce a native and unified approach to the coordination of radio access mechanisms in multi-radio access technology networks for efficient data delivery in mobile networks. This leads to a design of a unified traffic steering framework, aiming at the orchestration of traffic steering related features for optimal radio resource utilization. Load-based radio access network coordination is presented and accompanied by illustrative examples to visualize how the multitude of LTE-Advanced Pro features can be handled in a holistic manner. Based on the proposed solution, potential evolution directions of mobile networks are discussed as an initial step toward the standardization of 5G.

INTRODUCTION

Network densification techniques, new spectrum access schemes, and the increasing availability of radio access solutions geared toward spectral efficiency improvements are all enlarging the flexibility of the radio resource management (RRM) landscape. However, RRM flexibility enhancements yield an increase in complexity.

One example scenario addressing the spectrum access mechanism selection problem is presented to illustrate this problem: To cope with mobile data traffic growth, one approach may be to exploit carrier aggregation (CA) using macro sites acting as mobility anchors for moving user equipments (UEs), supported by dual connectivity (DC). Alternatively, small cell (SC) layers can offload time- and location-specific traffic, exploiting licensed or unlicensed spectrum bands. In the latter case, a licensed assisted access (LAA) [1] scheme may be more efficient for outdoor users due to the availability of interference management mechanisms. For the indoor environment however, an LTE-WLAN aggregation (LWA) [2] scheme might serve better due to the ease of WiFi access points deployment. In the case of SCs using the time-division duplex (TDD) spectrum, enhanced interference management and traffic adaptation (eIMTA) allows LTE frame adaptation to instantaneous uplink/downlink (UL/DL) traffic changes. Furthermore, multi-radio access technology (MRAT) joint coordination [3] includes mechanisms to manage the shared spectrum in a dynamic manner, for example, by opportunistically refarming the underutilized 2G spectrum and assigning it to LTE. Taking into account the selected alternatives mentioned above, it is far from obvious how to coordinate these solutions. This leads to the conclusion that to steer the traffic efficiently, a proper orchestration mechanism for radio access network (RAN) features is required.

To handle non-uniform data demand, traffic steering (TS) is considered one of the most important functions in current mobile networks, being responsible for routing user data traffic via the most suitable radio resources (RRs). In heterogeneous networks (HetNets) with a multilayer, multi-RAT, multicarrier landscape, candidate cell(s) assignment is not obvious [4], and the selection of serving cell(s) based on the best signal-to-interference-plus-noise ratio (SINR) is not always optimal. Additionally, mobile network operators (MNOs) may set different policies for TS operation. For example, TS may perform load balancing across carriers during the day and push the traffic to the coverage frequency layer for energy saving purposes during the night. In general, UE-specific cell selection is a multidimensional optimization problem. To properly assign radio links to users, a TS engine needs to consider a set of inputs (radio conditions, UE capabilities, available RATs and RAT-specific features, frequency bands and layers, cell load, quality of service (QoS) requirements, network and UE power consumption, etc.).

Selected aspects of this problem have already been addressed in prior work. The METIS II project has identified the need for a holistic approach in the "agile RRM" proposal, covering a multitude of use cases [5], while the COHERENT project addressed the Central RRM Coordinator framework [6]. These concepts, however, focus solely on 5G, and do not cover aspects of current 4G networks and the already existing problems. A coordination framework for HetNets was proposed in [7], but addressed only a selected number of functionalities, including SC

The authors are with Huawei Technologies Sweden AB.

0163-6804/16/$25.00 © 2016 IEEE

sleeping and cooperative transmission mechanisms. There are more references in the literature on this topic; however, due to the character of this article, we limit ourselves to provide the selected ones.

To address the complexity of the described problem of feature coordination and to address the limitations of prior work, this article proposes a holistic, scalable, and extendable unified traffic steering (UTS) framework. It covers Third Generation Partnership Project (3GPP) feature evolution and provides forward compatibility toward 5G. The design of the UTS focuses on the available spectrum access mechanisms and their orchestration. It also considers scenario-specific feature prioritization and ranking, aiming at optimal radio resources utilization. *UTS logic* serves as an engine, coordinating TS mechanisms with respect to varying mobile data load across a variety of scenarios, and provides system enhancements in terms of network coverage and capacity.

This UTS approach helps to avoid a situation where the handling of a multitude of fragmented solutions results in network instabilities or contradicting actions among individual RRM features. A UTS concept enables the integration of new features, additional spectrum bands, or novel spectrum licensing schemes. It was also considered for ultra dense networks (UDNs), as an evolution of HetNet scenarios moving toward 5G standardization.

The remainder of this article is organized as follows. After an overview of recent 3GPP developments in the area of TS related features, a UTS framework is introduced, and load-based feature coordination is discussed. Finally, there is a discussion of potential evolution directions of the LTE standard.

LTE-ADVANCED PRO FEATURES OVERVIEW

Recently, a new branding of the 3GPP standardization work has been announced, called LTE-Advanced Pro. It reflects further the evolution of the LTE-Advanced (LTE-A) standard, and maps to Release 13 and beyond [8]. Selected LTE-A Pro functionalities related to TS are highlighted below.[1]

LTE Carrier Aggregation: Evolution is continued, enabling spectrum aggregation beyond five component carriers (CCs). It enables so-called massive CA by allowing up to 32 CCs. It utilizes the physical uplink control channel (PUCCH) on a secondary cell (SCell) for UEs supporting UL CA. PUCCH is configured simultaneously on a primary cell (PCell) and on one SCell. Similar to CA, DC addresses user throughput enhancements. It focuses on mobile UEs, and aims at mobility robustness and signaling optimization in HetNet scenarios to avoid frequent handovers (HOs).

Multicarrier Load Distribution: Deals with load balancing solutions among LTE carriers in IDLE and CONNECTED mode, aiming at efficient spectrum resource utilization, avoiding unnecessary load-triggered handovers or redirections, thus decreasing signaling load and HO failure rates.

Unlicensed Spectrum Usage: It is considered in 3GPP Release 13 in the following ways:

- *LTE-WLAN radio level integration* (also known as LWA) enhances the higher-layer-based interworking mechanisms already specified for WLAN offload. It defines LTE and WiFi aggregation on the radio level for real-time channel and load-aware common RRM. The aim is to provide capacity and quality of experience (QoE) improvements in a HetNet environment. Scenarios cover non-co-located (DC-based) or co-located (CA-based) utilization of carrier WiFi serving as a secondary link.
- *Licensed assisted access* (LAA) covers CA-based scenarios, where a primary cell (PCell, i.e., the mobility and signaling anchor) operates in a licensed band, while at least one secondary cell (SCell) is using the unlicensed 5 GHz spectrum. It considers "fair coexistence" with WiFi, for example, using a listen-before-talk (LBT) mechanism.
- To enable a common LTE-WLAN RRM, the WLAN management work was triggered covering carrier WiFi performance monitoring and WLAN key performance indicators (KPIs) exchange with 3GPP RAN.

MRAT Joint Coordination: All of the above schemes fall under the MRAT joint coordination framework, which considers 3GPP RATs and carrier WiFi access. It aims at radio-resource-efficient MRAT management, reducing the configuration burden and signaling overhead caused by mobility management or load information exchange.

Licensed Shared Access (LSA): A new spectrum licensing method allows spectrum owners to share their resources with other MNOs [9]. An LSA advantage over the regular spectrum sharing scheme is QoS support, for example, by the use of a spectrum owner's defined protection or exclusion zones for shared bands. LSA work analyzes the support of static and semi-static sharing in 2.3–2.4 GHz band.

RAN Sharing Enhancements (RSE): Dynamic on-demand capacity negotiations and load balancing schemes for sharing scenarios are considered in [10]. The allocation of shared RAN resources is based on the proportion of assigned RAN usage for each of the participating MNOs.

SON for Adaptive Antenna System (AAS): Release 13 introduces a network densification technique, called self-organizing networks (SONs) for AAS. It allows automated cell shaping by splitting/merging in vertical or horizontal dimensions, as well as by beamforming. Cell shaping is triggered based on the cell load or interference conditions.

Table 1 covers details of the above features, gathering their advantages, limitations, and usage requirements.

UNIFIED TRAFFIC STEERING FRAMEWORK

Based on the above analysis of TS related features and their relations, a UTS framework is proposed (Fig. 1). It aims to manage the available spectrum resources with the available features taking into account varying (time-, location-specific) load conditions and MNO-specific TS policies. The goal of the UTS framework is to efficiently deliver capacity when and where needed to fulfill QoS requirements and to save energy in the network when and where possible.

A *mobile devices* entity collects the radio mea-

An individual UTS logic instance manages a single access node or a cluster of nodes. The decisions invoke standard RRC procedures for CONNECTED and IDLE mode actions or parameter settings for example, the assignment of new component carriers in CA.

---

[1] Detailed descriptions of all remaining Release 13 features are out of the scope of this article. A Release 13 features overview is covered in [11].

Feature	Input to/output from TS	Advantages	Disadvantages	Usage requirements	Use case
Carrier aggregation with enhancements beyond five component carriers	**Input:** Available CCs, load per CC **Output:** SCell assign/add/release, CC management, PCell assignment	• Improved UE throughput • Scheduling flexibility (cross-carrier) • No need for load balancing (LB) – CA scheduling distributes load among CCs	• Increased scheduling complexity • Increased RF complexity	• Multiple carriers • UE support	• Assign all CA-capable UEs with multiple CCs • Use cross-carrier scheduling for HetNet • Assign random PCell to each UE • Alternatives: DC, LAA, LWA
Dual connectivity	**Input:** User context, cell load, UE measurements **Output:** Secondary cell group (SCG) management	• Increased UE throughput • Improved mobility for HetNet • Improved cell edge performance • Decreased HO burden and LB	• Increased UE complexity • Increased scheduling complexity (split-bearer scheduler)	• Multiple carriers • UE support for handling two independent links • "Split bearer" scheduler	• Assign all DC-capable UEs with DC • Switch stationary users to be served by small cells • Assign DC to moving UEs or to cell-edge UEs (CEU) • Alternatives: CA, LWA, LAA, CoMP
eIMTA	**Input:** Cell load for DL/UL balance, co-channel interference measurements **Output:** DL/UL TDD frame configuration	• Improved RR utilization • Dynamic local traffic adaptation	• Extra signaling • Cell clustering required • Need for accurate interference measurements • Limited by HARQ signaling timing • Only for TDD SC	• UE support • Cell clustering mechanism	• More traffic in UL/more subframes for UL; • More traffic in DL/more subframes for DL • Alternatives: MLB, SON for AAS, eICIC
Licensed assisted access	**Input:** Available CCs, load per CC **Output:** CC management messages, unlicensed SCell assign/release	• Unlicensed spectrum usage enabled • Improved ROI for MNO	• Coexistence and access "fairness" with WiFi (e.g. LBT) • Scheduling complexity (second link management) • RF complexity • No QoS support • Lower performance than licensed carrier	• UE support • Unlicensed carrier available • CA capability as prerequisite	• Assign all LAA-capable UEs with unlicensed carrier if in coverage • Assign LAA if load on licensed is high • Apply LAA for outdoor scenarios • Alternatives: LWA, DC, CA
LTE-WLAN radio level integration	**Input:** Available SSIDs, cell loads, WiFi measurements **Output:** WiFi cell add/release, LWA management	• Enables usage of WiFi in a network-controlled manner • Improved ROI for MNOs	• Increased scheduling complexity (split bearer scheduler or CA scheduler) • No QoS support	• UE support (link aggregation) • Carrier WiFi only • "Split bearer" scheduler • Xw interface	• Assign all LWA capable UEs with LWA • Switch stationary indoor UE to the available WiFi • Apply LWA, if load in licensed carrier is high • Alternatives: LAA, DC, CA
Multi-carrier load distribution in LTE (MC load distribution)	**Input:** cell load, available carriers **Output:** UE IDLE mode camping info/reselection parameters	• Proactive UE distribution control among layers • Adds new, accurate measurements for reselection/cell camping/HO	• New UE measurements required	• New signaling • UE support (new measurements and procedures) • Need alignment with CONNECTED mode LB	• Distribute UEs among carriers (equally or proportionally) • Move cell-center users (CCU) to higher layers and CEUs to lower layers • Alternatives: MLB
Multi-RAT joint coordination	**Input:** traffic load per RAT, UE measurements (MRAT) **Output:** user steering decisions (HO and reselection between RATs)	• Improved RR usage efficiency and spectral efficiency by joint RR coordination • Guaranteed QoE across RATs • Improved network capacity by steering UEs to appropriate RAT, including carrier WiFi	• Large change in network management implementation	• Interfaces from different RATs • Measurement collection and processing among RATs • Possibly new KPIs	• UE class steering policy: steer video users to LTE, distribute among available layers (e.g. CA); steer voice users to GSM; • Steer stationary indoor users to carrier WiFi; • Steer stationary outdoor users to SCs (e.g. DC, LAA) • Alternatives: MLB
RAN sharing enhancements	**Input:** requested RR, required on-demand capacity, MNO priorities, RR usage monitoring **Output:** resource availability	• Dynamic RR usage optimization among MNOs in shared scenarios	• Required coordination with load optimization and interference management features (i.e., need updates on the actual RR availability)	• Shared RAN scenario • Accounting info per participating MNO • Inter-MNO signaling with respect to shared resources and required capacity	• Under high load, request more on-demand RR from the shared spectrum pool • If spectrum not used, release RR to the shared pool
SON for AAS	**Input:** cell load info (CEU/CCU ratio) interference measurements **Output:** cell split/merge	• Automatic mechanism for load triggered network densification (macro scenario)	• Proper coordination with conflicting features required (ICIC, MRO, MLB, CCO)	• AAS antennas • Proper measurements needed • Coordination with other features	• If load is high in the inner part of cell, trigger split cell • If interference too high to edge users, trigger merge cell • Alternatives: MLB, ICIC
WLAN management	**Input:** WLAN measurement configuration **Output:** QoS, load info, WLAN measurements	• Tight interworking enabled (3GPP/ WiFi) • Optimized TS decisions (enabler for LWA and offloading to standalone WiFi)	• Additional signaling • UE measurements needed	• UE support (additional measurements and reporting) • Interfaces between WLAN and 3GPP (Xw)	• Measure traffic load, resource availability, QoS with certain granularity and provide to eNB
License shared access (LSA)	**Input:** available transmit slots and spectrum portions **Output:** usage of shared spectrum (e.g. transfer UEs to LSA band)	• Additional spectrum resources enabled in time and location specific manner (e.g. request additional spectrum during busy hour) • QoS support on shared spectrum	• Increased RR coordination complexity • Implementation changes	• Spectrum broker entity and new interfaces • Coordination with cell reselection, HO, DC, CA, MLB to dynamically utilize LSA band	• Utilize available features (e.g. CA, DC, MLB) using additional spectrum bands (e.g. add new SCell from LSA band to the CA) • Alternatives: LAA, LWA (unlicensed schemes)

Table 1. Overview of selected Release 13 traffic steering features [11].

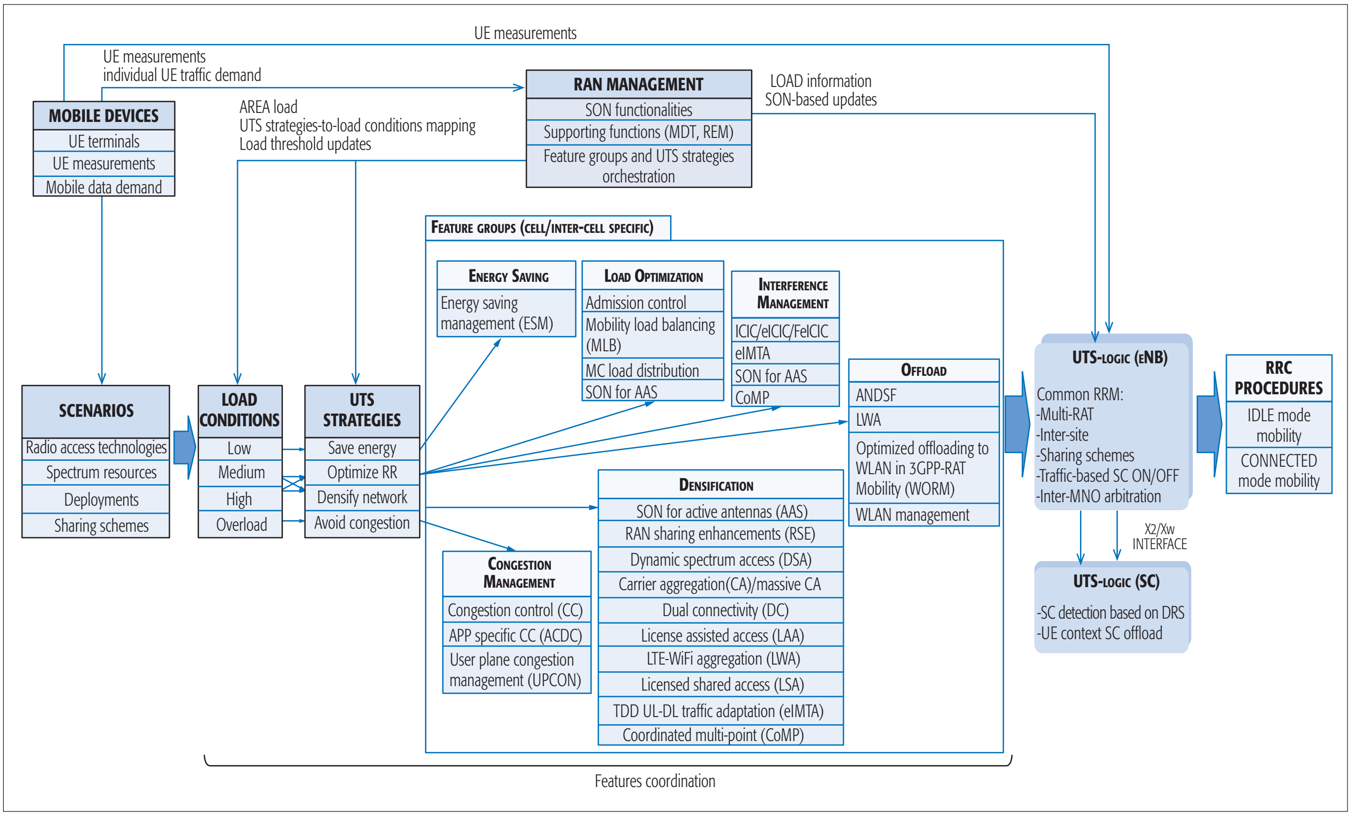


Figure 1. A unified traffic steering framework presenting the orchestration of a complex features landscape.

surements information and provides it for a TS decision and evaluation within SON functionalities handled by the *RAN management* block. Mobile data demand is considered in a scenario-specific manner within the *scenarios* entity (e.g., incorporating available RATs, spectrum resources, and sharing schemes). It enables the definition of the load conditions, which are then mapped to the MNO-specific UTS strategies (e.g., improve user experience by densification features and optimize radio resources usage). The feature coordination goal is to orchestrate the available features, providing input to *UTS logic* for decision processing and optimization.

Aiming at efficient feature-to-UTS strategy mapping, the following feature groups were identified.

Energy saving (ES): Switch off underutilized cells; reconfigure neighbor cells to compensate coverage:
- *Macro ES*: Coverage or capacity layer energy saving management (ESM) with coverage compensation
- *Small cell ES*: sleeping mechanisms with discovery reference signals (DRS) utilization

Load optimization (LOpt): Methods for IDLE and CONNECTED mode load balancing:
- *Macro LOpt*: Admission control, mobility load balancing (MLB), MC load distribution
- *Small cell LOpt*: Admission control, eIMTA, enhanced inter-cell interference coordination (eICIC)

Interference management (IM): Methods for managing co-channel interference:
- *Macro IM*: Inter-cell interference coordination, coordinated multipoint (CoMP)
- *Small Cell IM*: eICIC, eIMTA, CoMP

Densification: Utilization of multiple carriers, cell splitting or utilization of SCs:
- *Macro*: CA, SON for AAS.
- *Small cell*: CA, DC, LAA.

Offload: Utilization of WiFi nodes to offload LTE cells including:
- LWA, ANDSF-based WiFi offload, WLAN management, optimized offloading to WLAN in 3GPP RAT mobility (WORM)

Congestion management: Traffic class prioritization including:
- *Application-specific congestion control for data communication* (ACDC); *user plane congestion management* (UPCON).

UTS logic is the central entity of the framework. It makes coordinated decisions triggering specific features depending on the load conditions. An individual *UTS logic* instance manages a single access node or a cluster of nodes. The decisions invoke standard RRC procedures for CONNECTED and IDLE mode actions or parameter settings (e.g., the assignment of new component carriers in CA). *UTS logic* is fed with the UE capabilities and measurements, radio link conditions, available radio resources, RATs, and the nodes' capabilities, as well as inputs from the *RAN management* entity (e.g., cell-specific load or SON updates).

RAN management automates the UTS framework to flexibly assign a UTS strategy to the current network deployment in a specific area, for example, depending on the availability of SCs or sharing schemes. It handles the following functionalities.

Feature group orchestration, responsible for:
- Feature priority updates, when network

An individual instance of the UTS logic coordinates usage of the spectrum resources, RATs and small cells under single Macro Cell Coverage Area (MCCA), which defines the geographical area within the Macro cell coverage (for nominal antenna configuration), serving as a coverage layer.

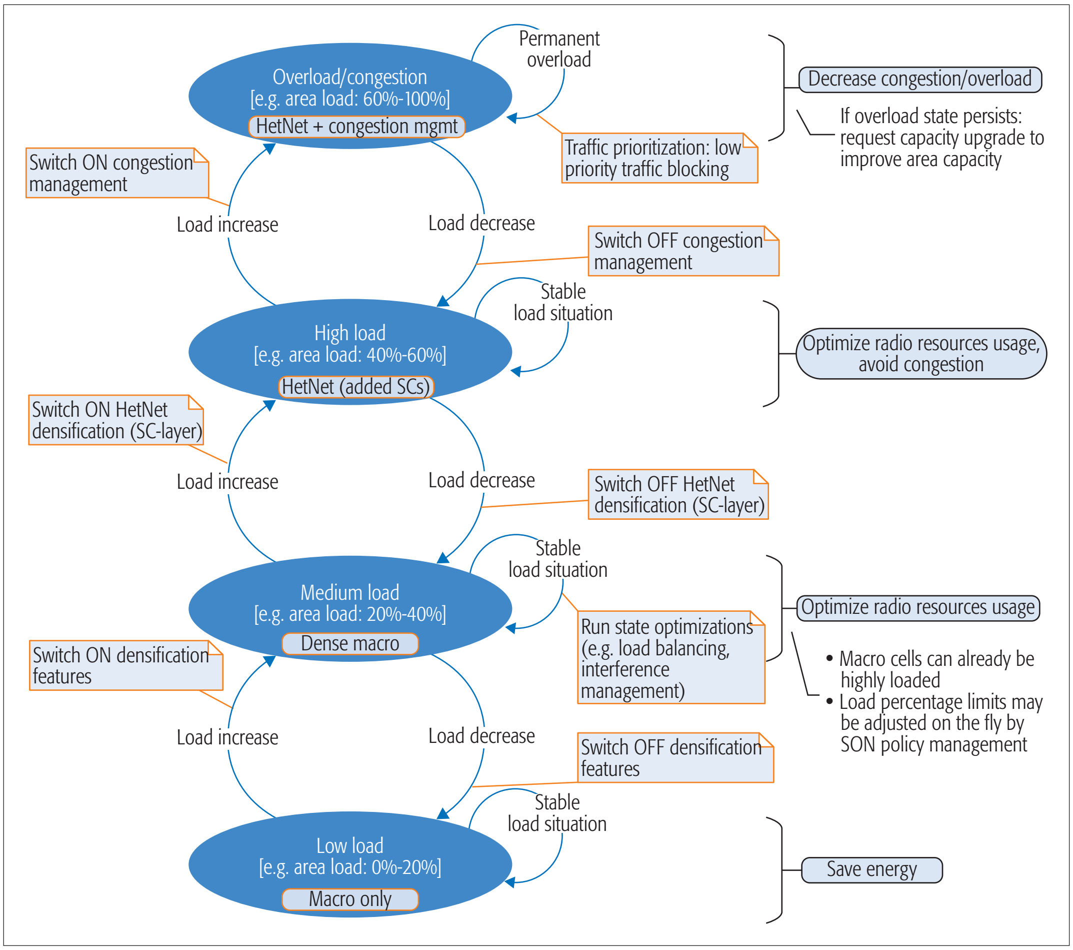


Figure 2. Load-based state diagram.

capacity is upgraded (e.g., more carriers are available)

- Feature ranking updates promoting more aggressive offloading when a network evolves (e.g., more SCs are deployed)
- Suggesting network capacity upgrades in case long-term congestion is identified in a certain area

SON features include:

- Standard self-optimization features (e.g., automatic neighbor relationship, mobility robustness optimization, and capacity and coverage optimization). These features are always enabled to optimize radio resource usage and to improve network KPIs.
- Standard self-configuration and self-healing mechanisms not directly related to TS.
- Measurements and QoS information supported by minimization of drive tests (MDT) and radio environment maps (REMs) presenting radio/interference conditions and user perception for proper traffic steering policies updates.

UTS FRAMEWORK IMPLEMENTATION CONSIDERATIONS

Based on the above description, the key differentiators characterizing the proposed UTS framework are as follows.

Implementation aspects of a UTS framework:

- *UTS logic* is distributed among cooperating instances that are exchanging information (e.g., load, interference coordination messages, neighbor relation updates).
- *UTS logic* instances implemented per macro eNB nodes and X2/Xw interfaces are used for inter-site coordination.
- SCs act as slave nodes under master macro eNB management, controlling SCs within their coverage layers (to enable fast ON/OFF mechanisms and control plane/user plane split).
- *UTS logic* instances with a limited set of functionalities are located in SCs for the purpose of low load detection and switch-off triggering.
- A logically centralized *RAN management* entity orchestrates the implementation of distributed *UTS logic* instances.

The timescale of the UTS operations has to address dynamic load variations in HetNet scenarios.

- Traffic pattern collections and big data predictions of load patterns in the network performed within the *RAN management* entity are used for the dynamic parameterization of the UTS action triggers in the time domain (directly impacting the amount of the signaling overhead).
- The UTS framework does not imply per transmission time interval (TTI) granularity, in order to avoid excessive signaling overhead and potential instability of the system. The measurements, based on which actions are taken, should

be low-pass filtered, that is, ranging from several tens of milliseconds up to seconds.

Signaling overhead of the proposed framework.

- Radio interface signaling includes typical LTE compliant radio resource control (RRC) (re)configuration messages to assign radio links to UEs. Thus, it does not introduce additional signaling overhead compared to a system without a UTS framework. The difference comes from the fact that the UTS-based operation relies on the overall coordinated picture of radio conditions within the cluster, obtained from the UE measurements, which are further heavily low-pass filtered by the *UTS logic* to limit the overhead. The signaling overhead results from the actual algorithm running within the framework to make decisions on radio link management (e.g., change/add/delete link).
- Coordination signaling results from the interworking between *UTS logic* entities, as well as from required features interaction. The interaction includes inputs to and action triggers from the *UTS logic* (e.g., cell ON/OFF messages, eICIC updates for interference management among Small Cells, etc.). As the UTS framework operates on Layer-3 level, non-real time signaling is required and thus does not provide excessive RAN overhead.

***UTS logic* operates using layer 3 reconfigurations on the radio interface; however, it interacts with the physical (PHY) and medium access control (MAC) layers in the following ways.**

- PHY layer: In order to efficiently utilize sleep mode in the case of SCs used as capacity boosters serving the user plane layer only, a "light" radio frame (e.g., lean carrier) with DRSs is required. Additionally, the interference conditions are changed by adaptive operation using eIMTA or eICIC coordination mechanisms.
- MAC layer: UTS impacts scheduling decisions with the use of eIMTA, eICIC, or CoMP features. On the other hand, *UTS logic* decisions triggering SC OFF or eIMTA frame adaptations require measurements provided by the MAC layer (e.g., buffer reports, throughput, interference conditions).

LOAD-BASED FEATURE COORDINATION WITHIN THE UTS FRAMEWORK

Individual *UTS logic* controls feature operation via actions taken upon the load changes within an access node or within a cluster of nodes. Load changes are mapped to load-based state machine diagrams (Fig. 2). Load-state definitions, state transitions, and related triggers depend on the available feature set, UTS strategies, as well as traffic load characteristics. The feature coordination approach is as follows:

- *Upon each load state transition*: Features are switched on/off.
- *Within each predefined load state*: All enabled features are operational with internal updates/optimizations (including conflict resolution among features).

UTS strategies are defined per load state and selected by the MNO; for example, in case of medium load state, macro network densification is triggered. Proposed UTS strategies[2] for the load states (Fig. 2) are:

- Low load: *save energy*: Utilize the macro layer only; switch OFF certain carriers/cells to save energy during low traffic periods.
- Medium load: *densify network, increase macro layer capacity*: Add carriers, split cells, and enable CA. Improve resources utilization and interference across layers: macro IM, macro LOpt, and AAS.
- High load: *densify network, enable SC layer*: Switch ON SCs in hotspots; enable LAA, LWA, and DC. Improve resources utilization and interference across layers: enable SC interference management and SC load optimization.
- Overload: *decrease congestion*: Utilize congestion management features upon overload detection.

An individual instance of the *UTS logic* coordinates usage of the spectrum resources, RATs, and SCs under a single *macrocell coverage area* (MCCA), which defines the geographical area within the macrocell coverage (for nominal antenna configuration), serving as a coverage layer. The following UTS metrics are defined per MCCA.

Area capacity (ACap): is defined as 100 percent of the network capacity available over the MCCA coverage layer per macrocell. It includes the capacity of all licensed frequency bands[3] and all layers (including SCs). ACap is a scalar value defined per MCCA (e.g., ACap = 5 Gb/s/MCCA) that can be obtained in various ways:

- Channel-quality-indicator-based: estimation using spectral efficiency statistics within an MCCA (i.e., weighted average of individual MCS usage over the area)
- MDT-based: UE-report-based estimation; reports from high-load periods
- REM-based: estimation using the location-specific dynamic interference maps

Area load (AL): is defined as the utilization of available ACap. AL = 100 percent is the ACap for a particular MCCA.

Area load threshold: is defined as a percentage of the ACap, used to trigger load-state transition, which in turn enables proper feature groups. Area load thresholds are subject to optimization per MCCA by SON mechanisms within the *RAN management* entity. Statistics of the RAN usage per MCCA are collected and stored within *RAN management* for thresholds tuning purposes. Based on potential long-term overload-state duration, network capacity extensions and upgrades could be decided by the MNO.

Statistics of the RAN usage per MCCA are collected and stored within RAN Management, for thresholds tuning purposes. Based on potential long-term overload-state duration, network capacity extensions and upgrades could be decided by the MNO.

FEATURES COORDINATION EXAMPLE

UTS strategies are defined, or selected from the predefined set, by the MNO. These strategies should reflect the scenario-specific traffic characteristics. To visualize the range of scenarios and features covered by the UTS framework an example use case is presented (Fig. 3).

The proposed scenario consists of two areas: a dense urban (DU) area and a suburban (SUB) area. The traffic density and daily traffic pattern differs between these areas. In the case of DU areas, traffic is in bursts and condensed within a limited area, while within SUB areas traffic variations are smoother and distributed over a wider

---

[2] In the UTS proposal, triggering events are considered to be load-based. Other triggers (e.g., cell outage, handover KPI changes) can be evaluated using the same methodology, and increasing the framework scope.

[3] Estimated capacity of the unlicensed spectrum may introduce high uncertainty to the available ACap due to the lack of control over the interference level.

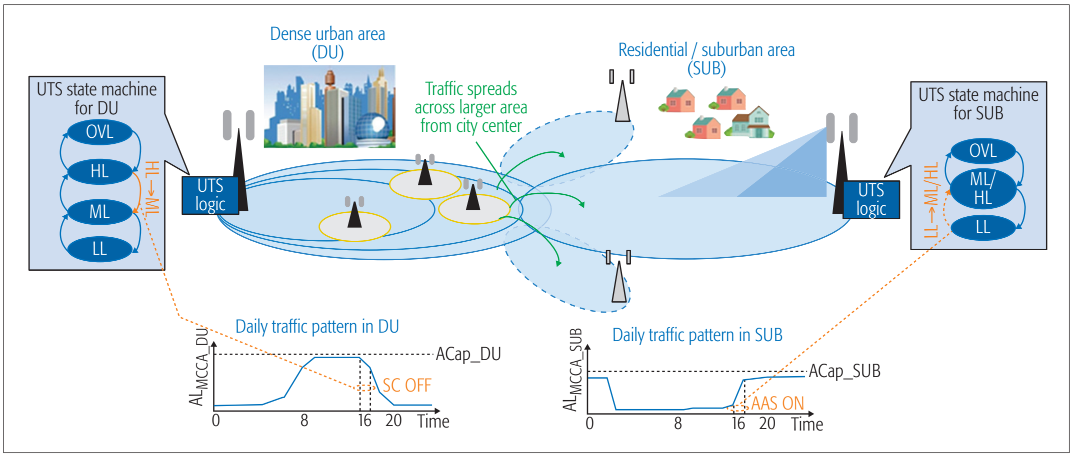


Figure 3. Evaluation scenario.

area. The number of load states for DU areas and SUB areas also differs due to different traffic pattern characteristics, different UTS strategies, and different features used in those areas (Table 2). The network in a DU area consists of multi-band macro layer and SC layer, including various RATs (i.e., LTE-FDD, LTE-TDD, LTE-U, and WiFi). A suburban area is served by a multi-band macro layer with AAS antennas for densification purposes. Each area is covered by a separate MCCA. There are two *UTS logic* entities involved, handling and coordinating features and radio resource usage within individual MCCAs.

The load state transition is analyzed, in which the traffic shifts from DU areas toward a SUB area, modeling a late afternoon situation when most of the traffic is moving from the city center (business area) to residential areas. *Area load* (AL) for *MCCA_DU* peaks during daytime, whereas the AL for *MCCA_SUB* peaks during the evening/night. It is assumed that the AL in *MCCA_DU* is much higher than the AL in *MCCA_SUB*, justified by the fact that the traffic concentrated over a relatively small DU area during working hours (covered by a single MCCA) spreads across multiple surrounding residential areas (covered by multiple MCCAs).

UTS logic in DU areas detects a load decrease from high load to medium load. According to the UTS strategy defined for this state transition (Fig. 2), SCs are switched off whenever possible, for

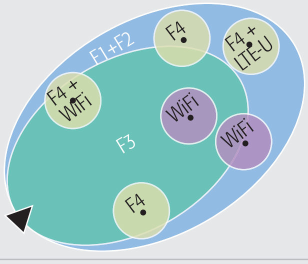
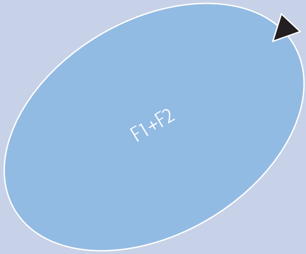
Scenario	Dense urban (DU) – City center	Suburban (SUB) – residential area
Network configuration	**MCCA structure:** Dense macro and SCs in hotspot areas; **Macrocell carriers:** F1 (FDD), F2 (FDD), F3 (FDD) **SC carriers:** F4 (TDD), F5 (unlicensed) 	**MCCA structure:** regular macro **Macro cells carriers:** F1 (FDD), F2 (FDD) **SCs carriers:** NO 
Load distribution and load variations	• Very low during night, increasing in morning, rush hours during daytime, decreasing in the evening • Hotspots, highly non-uniform, bursty traffic	• Constant low traffic during the day, more traffic during the evening and early night • Regular, uniform
Load states	1) LOW, 2) MEDIUM, 3) HIGH, 4) OVERLOAD	1) LOW, 2) MEDIUM/HIGH, 3) OVERLOAD
Carrier WiFi/open access WiFi	Yes (e.g., shopping malls, municipality deployed, public transportation with onboard Wifi nodes)	No
UE mobility	Stationary, pedestrian, vehicular: connected cars/Mi-Fi (30km/h), commuters	Stationary, pedestrian, vehicular: connected cars/MiFi (50 km/h+)
Main TS features	LAA, LWA, CA, DC, eIMTA, RSE (4G capacity sharing)	AAS, CA
Scenario challenges	• Challenging interference due to uncoordinated customer premises SC deployment (HeNB) • High capacity demand • Complex coordination among features/RATs • Site acquisition problems in mature markets	• Coverage holes (especially for 4G) • Backhaul availability

Table 2. Scenario specification.

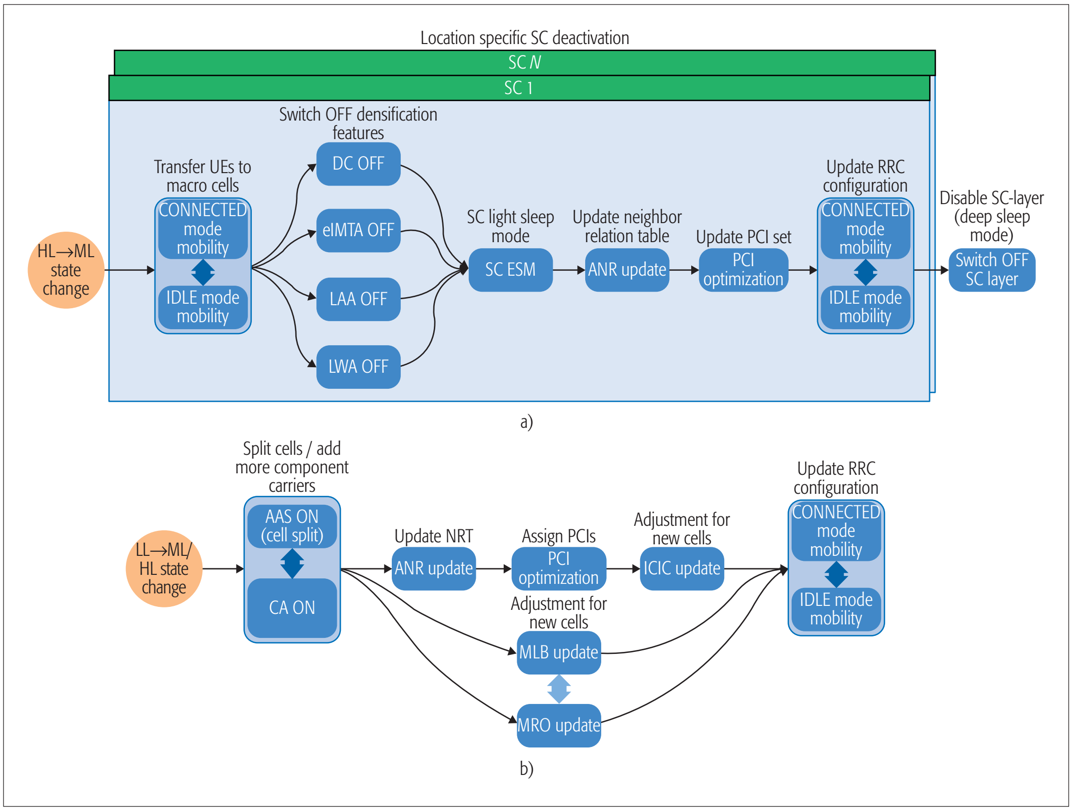


Figure 4. a) Dense urban state transition (high load to medium load); b) suburban state transition (low load to medium/high load).

UTS logic in dense urban areas detects a load decrease from high-load to medium-load. According to the UTS Strategy defined for this state transition, small cells are switched off whenever possible, for energy saving and interference management purposes.

energy saving and interference management purposes. The following processing is performed by the *UTS logic* per individual SC node (Fig. 4a):

- UEs are moved from SCs to macro carriers utilizing handover/redirection commands, and CONNECTED and IDLE mode parameters changes are updated.
- *Interference Management*, *Load Optimization* and *Offload* features are switched-off.
- ESM is activated for those SCs, triggering sleep mode (each SC is individually switched OFF in adaptive fashion when the traffic is low within that particular SC).
- The neighbor relation table (NRT) and PCIs are adjusted according to the new neighbor situation. RRC protocol updates measurements and cell configurations with respect to changes from ANR, PCI optimization, and ESM functions.
- When all SC are in sleep mode, the whole SC layer is switched off (i.e., SC switch to deep sleep mode, thus no longer sending DRS signals).

As the traffic shifts away from the DU area, the *UTS logic* in the SUB area detects the load increase from low load to medium load. The UTS strategy here is to densify the macro layer and increase the available capacity (Fig. 4b):

- *Carrier aggregation* with a second component carrier and SON for AAS enabling cell splitting is activated.
- Interference management and load optimization mechanisms are switched on and neighboring relations are updated accordingly:
  – MLB is adjusted to properly distribute traffic among new cells.
  – MRO settings are updated for cell splitting (note: MRO parameters can be optimized separately for split and non-split case, to avoid a constant need for adjustments each time the cell is split/merged).
  – MLB and MRO are coordinated, to avoid conflicts between them (e.g., MLB may want to move traffic toward one cell due to load imbalance, while at the same time, MRO may trigger an opposite action due to mobility performance issues).
  – NRT is updated by ANR, followed by PCI adjustment.
  – ICIC is updated to handle new interference scenario.
- RRC invokes CONNECTED/IDLE mobility procedures and parameter changes to reflect feature updates.

Summary and Future Evolution

Recent developments in the area of LTE evolution have been briefly covered in this article. Based on the observed "RRM landscape" complexity, the authors claim that traffic steering requires a unified and future-proof approach to achieve radio-resource-efficient networks. The UTS framework is proposed as a solution to orchestrate a multitude of TS-related RAN functionalities and mapping instantaneous mobile data demand to the available spectrum resources. Non-uniform traffic demand, stemming from HetNet scenarios, are evolving toward more complex cases along with the introduction of novel services like vehicular communication or a wide range of IoT applications. To cope with such variable mobile data demands in future net-

LTE evolution will be continued in parallel to NR/5G development to address service continuity and fallback solutions for mobility management. In that context the UTS framework is continuously evolving to cover both the new aspects of 5G as well as the support of interworking between evolved LTE and NG RAT.

works, it is claimed that the access node capabilities, their locations, and the resulting density should also be deployed in a non-uniform manner. This requires an adaptive unified traffic steering framework, posing requirements on software defined network based deployments, to address the limitations of static network planning. UTS assumes an awareness of the traffic demand and the ability to optimize its power consumption within HetNet networks. This requires radio resource coordination on multiple levels: inter-RAT, inter-band, inter-site, inter-layer. *UTS logic*, the central node in the proposed framework, is considered to serve as an engine to coordinate TS mechanisms with respect to scenarios, feature groups, MNO policies, traffic steering strategies, and load conditions. In conclusion, the expected UTS framework benefits are:

- Unified orchestration of traffic steering features to decrease the potential instability of network operation
- Optimized usage of small cells, enabling energy efficiency by the introduction of modified RRM states with two-stage sleep mechanisms
- Forward compatibility and easy integration of new RAT and related features within the UTS framework
- Consistent and unified decisions on radio links selection to serve user traffic, and to enable mobility support for multi-connectivity in future networks.

The future-proof formulation of the UTS framework was one of the working assumptions during its design, to allow its utilization in multi-RAT scenarios moving toward 5G standardization. 5G is expected to address a mobile communications paradigm shift, shifting from the current network-centric approach toward user-centric concepts, such as novel performance metrics handling in the form of user focused key quality indicators or a cell-less network design approach, supported by native SON.

To address these developments beyond LTE-Advanced Pro, the 3GPP System Architecture and RAN groups have already started discussions on Release 14 requirements. From a UTS framework point of view, the relevant architectural aspects were discussed in the Study on Architecture for the Next Generation System, which is expected to define the requirements for a next generation (NG) system [12]. According to this study, NG RAN architecture should support:

- Tight interworking between new RAT and LTE, covering inter-RAT mobility as well as multi-RAT aggregation, for collocated and non-collocated deployments
- Multi-connectivity via multiple transmission points, for collocated and non-collocated deployments
- Separation of control plane signaling and user plane data from different sites
- Interfaces for inter-site scheduling coordination
- Network slicing, allowing different logical networks to utilize the same network infrastructure
- Harmonization of MAC and higher layers
- Network functions virtualization concepts in the context of RAN architecture

The currently discussed set of tools and mechanisms within the spectrum access domain is expected to further evolve in coming 3GPP releases. For example, the addition of millimeter-wave spectrum bands to the RRM toolset opens new opportunities for future mobile broadband applications. It also creates challenges due to demanding channels and network architecture implications requiring coordination with the legacy RATs for fallback solutions.

In parallel to the initial Release 14 studies, 3GPP RAN recently held a 5G RAN Workshop collecting contributions to the 5G vision proposals, requirements, and potential solutions. The discussion indicated that the 5G network service requirements will have to cover a much wider scope of applications than is being handled in current 4G networks, such as automotive, health, energy, or manufacturing applications. This leads to a new, non-backward-compatible RAT requirement to be introduced by 3GPP under the 5G umbrella. LTE evolution will be continued in parallel to NR/5G development to address service continuity and fallback solutions for mobility management. In that context, the UTS framework is continuously evolving to cover both the new aspects of 5G as well as the support of interworking between evolved LTE and NG RAT.

REFERENCES

[1] 3GPP TR36.889, "Feasibility Study on Licensed-Assisted Access to Unlicensed Spectrum," July 2015.

[2] 3GPP RP-150510, "LTE-WLAN Radio Level Integration and Interworking Enhancement," Mar. 2015.

[3] 3GPP TR 37.870, "Study on Multiple Radio Access Technology (Multi-RAT) Joint Coordination," June 2015.

[4] P. Fotiadis *et al.*, "On the Potentials of Traffic Steering in HetNet Deployments with Carrier Aggregation," *Proc. IEEE VTC*, Seoul, Korea, 2014

[5] O. Bulakci *et al.*, "Agile Resource Management for 5G: A METIS-II Perspective," *Proc. IEEE Conf. Standards for Commun. and Networking*, 2015, pp. 261–66.

[6] 3GPP RWS-150086, "COHERENT Vision on Software Defined Networks for 5G," 3GPP RAN 5G Wksp., Sept. 2015

[7] G. K. Tran *et al.*, "Dynamic Cell Activation and User Association for Green 5G Heterogeneous Cellular Networks," *Proc. IEEE PIMRC*, 2015, pp. 2364–68.

[8] D. Flore, "Initial Priorities for the Evolution of LTE in Release-13," 3GPP RAN, Sept. 2014.

[9] ETSI TS 103 154, "System Requirements for Operation of Mobile Broadband Systems in the 2300 MHz–2400 MHz band under Licensed Shared Access (LSA)," Oct. 2014

[10] 3GPP TR 22.852, "Study on Radio Access Network (RAN) Sharing Enhancements," Sept. 2014.

[11] 3GPP RP-151569, "Release 13 Analytical Vew Version Sept. 9th 2015," Sept. 2015.

[12] 3GPP RPA-160078, "Requirements on RAN Architecture," Jan. 2016.

BIOGRAPHIES

MARCIN DRYJANSKI (marcin.dryjanski@huawei.com) received his M.S. degree in telecommunications from the Poznan University of Technology in 2008. He served as senior R&D engineer and lead expert at IS-Wireless responsible for architecting of IS-Wireless' software solutions and providing expert level training on LTE/LTE-Advanced. He was a Work Package Leader in FP7-5GNOW project targeting 5G radio interface design. Currently he serves as a RAN system specialist at Huawei Technologies desiging algorithms and architecture toward 5G. He is an expert in PHY/MAC/RRM design and a co-author of numerous research papers targeting 5G RAN design.

MICHAL SZYDELKO [M] (michal.szydelko@ieee.org) received his M.S. degree in mobile telecommunication from Wroclaw University of Technology in 2005. He was involved in 3GPP RAN4 work as Nokia Siemens Networks delegate, taking responsibility for BS demodulation requirements for WCDMA and LTE/LTE-A. He was involved in the FP7-ICT SAPHYRE research project on spectrum sharing performance evaluation. He co-authored a number of research papers in the area of spectrum sharing and mobile networks evolution toward 5G. Currently he serves as a radio research engineer at Huawei Technologies Sweden AB, working on system architecture and radio networks evolution toward 5G and UDN networks.