Design of Cognitive Cycles in 5G Networks

Bego Blanco, Jose Fajardo, Fidel Liberal

Design of Cognitive Cycles in 5G networks

Bego Blanco, Jose Oscar Fajardo, and Fidel Liberal

School of Engineering of Bilbao, University of the Basque Country (UPV/EHU),
Alda. Urquijo S/N, 48012 Bilbao, Spain
`{bego.blanco,joseoscar.fajardo,fidel.liberal}@ehu.eus`

Abstract. Adding cognitive capabilities to the wireless networks makes it possible to leverage the control and management information used in the network operation to infer information about the local state and exploit it to improve the overall performance. This paper deals with the combined use of centralized and distributed cognitive cycles integrated at different planes in 5G networks: an integrated data plane, a unified control plane and a cross-layer management plane. This context-aware cognitive schema acts on the decision making modules depending on the monitored environment to prevent failures, balance the virtualized execution and get a global enhancement in the provision of mobile services. The multi-level cognitive cycle supports the interaction between the edge and the cloud blurring the line that separates two paradigms: centralized radio operation and mobile edge services.

Keywords: 5G, cognitive cycle, glocal, blurring edge

1 Introduction

The challenging requirements of 5G demand an evolution of the existing mobile network architecture. A defining characteristic of 4G networks is the coupling of data and control planes at the network level that interact with a separate service plane. This architecture forces to an adaptive operation: the service layer declares its requirements to the network layer; then the network accommodates the new service together with the other service requests and communicates the service plane the Quality of Service (QoS) levels that can be provided; finally, the service plane rearranges its requirements adapting them to the QoS offer. This iterative process is performed in isolation at both service and network planes without complete information of the other side, leading to non-optimal adaptation of services to network performance and of networks to service requirements.

In order to overcome these drawbacks, an innovative feature of 5G networks against the previous architectures is the clear decoupling of data and control planes and the integration of service management elements in the control plane, while service instantiations are allocated in the data plane. In this context, Software Defined Networking (SDN) provides the abstraction tool needed to split control and data planes in data forwarding. The control plane is now centralized and has the complete picture of the network, even considering different access technologies. Eventually,

this plane is capable of driving the different traffic flows along the most suitable paths and, thus, optimizing the use of the available resources in a service/network cross-layered way. At this point, context-aware cognitive techniques can become a useful tool to support this optimization process.

In this context, the data and control plane decoupling of 5G is aligned with the Mobile Edge Computing (MEC) principles. The application of MEC to 5G systems allows the physical separation of the planes, leaving the data plane close to the user in the network edge and uploading the centralized control plane to the cloud servers.

Another evolution step of 5G systems is provided by Network Function Virtualization (NFV) technology. NFV decouples the network functions from proprietary hardware appliances so they can run in software on a data center and can be instantiated in various locations in the network as required. Combining NFV with cloud computing concepts, a centralized orchestrator is responsible for the on-boarding of new network services and VNF packages together with the management of network service lifecycle. Again, at this stage, a second level of cognitive cycle can support the decision making process of the NFV Orchestrator (NFVO).

This paper is organized as follows: Section 2 makes a brief introduction of 5G overall architecture and its enabling technologies, i. e. SDN, MEC and NFV. Then, Section 3 reviews relevant literature about the integration of cognitive capabilities in network systems. Next, Section 4 shows an application proposal of a cloud-based multi-level cognitive cycle to the control and management planes of H2020 SESAME architecture [1]. Finally, Section 5 ends summarizing the main conclusions.

2   5G overview

Although the development of 5G is still in its early steps, the research effort in this area is oriented to achieve to main overall objectives: meet the requirements of the biggest mobile traffic growth ever known in a sustainable way, and provide a consistent end-to-end experience under diverse scenarios with ultra-high data rate, ultra-low latency and massive connections. Based on the analysis of these requirements, the International Telecommunication Union (ITU) [2] proposes the high-level network architecture depicted in Fig. 1.

An innovative feature of 5G systems against previous architectures is the clear separation of control and data plane functions, with open interfaces defined between them in according with SDN principles [3]. The decoupling of hardware and software functions of network elements in all network domains fosters a cost efficient deployment and upgrade possibilities. Other inherent benefits are real-time and on-demand network configuration and automated optimization, flexible and cost efficient network operation, maximization of utilization efficiency of available network resources and dynamic relocation of network resources, fully controlled by the operator. The disengagement of the network control logic and its centralization in an upper level provides a comprehensive view of the network state, so that the Network and Service Orchestration module is capable of configuring and managing the network service (NS)

lifecycle (including instantiation, maintenance and termination), while the data plane allocates the NS instances.
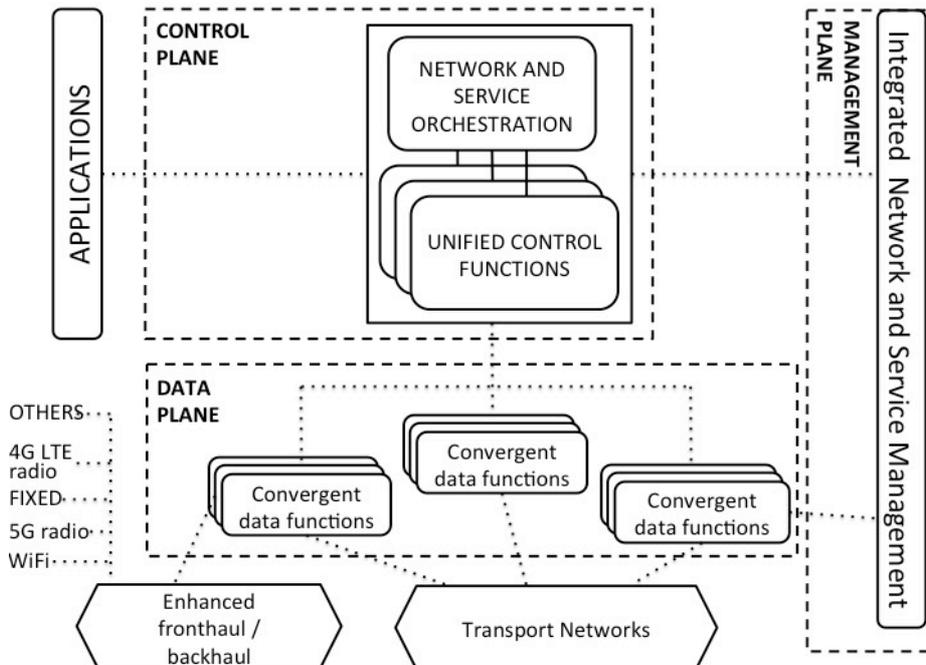

Fig. 1. High level network architecture of 5G based on [2].

This decoupling of control/management planes from the data plane is aligned with the concept of Mobile Edge Computing (MEC). MEC provides IT and cloud computing capabilities within the Radio Access Network (RAN) close to the end user [4]. Traditionally, all data traffic originated at the data centers is forwarded up to the mobile core network and down again to a base station which delivers the content to the mobile devices. In the MEC scenario, cloud servers take some or even all of the tasks originally performed in the data centre and eliminate the need of routing these data flows through the core network. This way, the RAN edge offers a low-latency high-bandwidth service environment as well as direct access to real-time radio network information that can be used by the upper plane to provide context-aware services.

Furthermore, the combination of SDN and MEC techniques over 5G architecture enables the application of Self Organizing Network (SON) principles to replace the classic manual configuration, post deployment optimization, and maintenance in cellular networks with self-configuration, self-optimization, and self-healing functionalities [5]. SON technology constitutes a fundamental change in the way networks are managed, bringing automation and dynamic, predictive resource allocation to the forefront.

Finally, the integration of a NFV framework takes another step forward in the optimization of the system performance. Fig. 2 shows the NFV architecture proposed by ETSI ISG NFV [6].
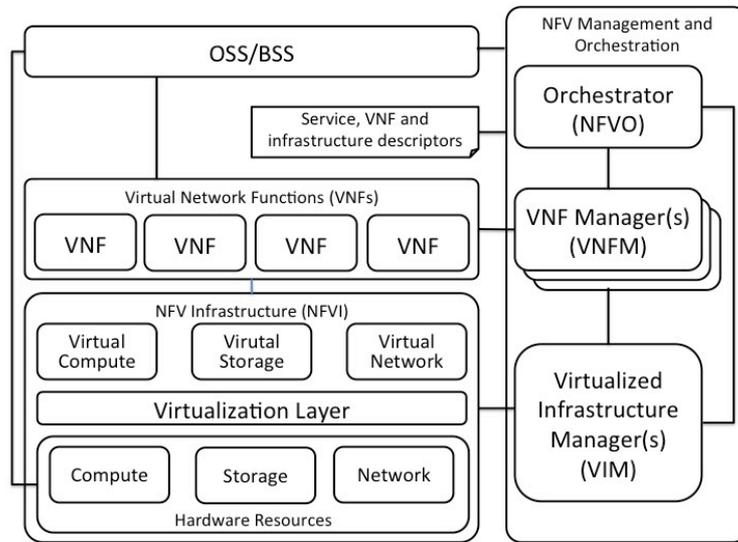
Fig. 2. ETSI NFV architectural framework based on [6].

The NFV concept envisages the implementation of Network Functions as software-only entities that run over the NFV Infrastructure (NFVI). As part of the NFV MANO, the NFV Orchestrator manages the incorporation of the Network Services through the chaining of VNFs and global resource arrangement. The VNF Manager oversees lifecycle management of VNF instances and coordinates the configuration and event reporting between NFVI and E/NMS. Finally, the Virtualized Infrastructure Manager (VIM) controls and manages the NFVI compute, storage, and network resources.

Next section analyzes the existing literature on cognitive cycles in order to insert intelligence into the different decision making elements in the network-service provisioning.

3   Cognitive networks

The philosophy of a cognitive network is a generalization of the renowned Cognitive Radio model [7]. This concept is proposed to make a more efficient use of the electromagnetic spectrum in wireless networks. The main idea consists in the variation of transmission and reception parameters according to the observed internal and external factors. Cognitive radio, built over software defined radio, is described as an

intelligent wireless communication system that is environment aware and learns from it, adapting to the statistical variations of a set of indicators.

The cognitive network [8] extrapolates the concept of cognitive radio to other levels of the communications reference model, minimizing the dependence on human intervention and adapting to the continuous changes of the network conditions and user requirements fast, precise and automatically. The key characteristic of this kind of network is the cognitive process that monitors the current state of the network, plans a future action, makes a decision and acts in consequence. With this objective, a cognitive network must identify the network conditions and predict future situations, constantly adapt to the dynamic state of the network, learn from previous experiences and balance the requirements of all the participants according to service agreements [9].

These characteristics lead to the cognitive cycle proposed in OOPDAL (Observe, Orient, Plan, Decide, Act, Learn) model [10] depicted in Fig. 3.
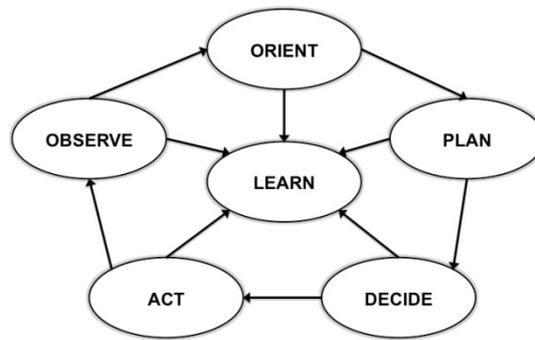
Fig. 3. OOPDAL cognitive cycle model based on [10].

The observation phase collects raw data to feed the orientation phase. The planning phase must identify the goals to achieve and arrange possible actions depending on the current situation to get closer to those goals. The decision phase assesses the available action plan and chooses the most appropriate. The action phase executes the selected plan, which may consist in a reconfiguration of parameters or a replacement of operative modules. Finally, the learning phase receives the feedback from all the previous phases and uses that experience to update the models employed in prevision, orientation, planning and decision processes.

On the other hand, Motorola-IBM proposes a simplified cognitive model within their FOCALE (Foundation-Observation -Comparison-Action-Learn-rEason) architecture for the autonomous network management [11]. Although this framework does not explicitly include a cognitive cycle, the description of its control cycles contains the necessary elements to recognize a cognitive process that involves monitoring, analysis, planning and execution tasks linked by a learning process. While it is a sim-

pler cycle than OOPDAL, it adds the novelty of associating the learning with business goals, something that is not considered in OOPDAL.

CME (Cognitive Management Entity) architecture [12] also suggests a 5 phase cognitive model (as in Fig. 4).
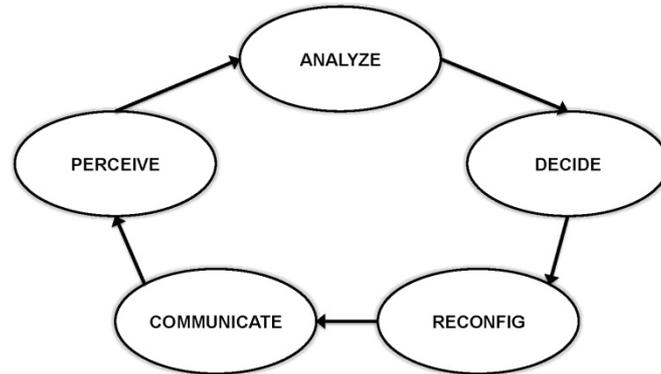

Fig. 4. Cognitive cycle of CME architecture based on [12].

The first phase is perception, responsible for collecting, measure, monitoring, pre-processing and approaching the context or environment information. Then, the analysis phase takes charge of the interpretation and abstraction, the consistence check, the prediction, the information fusion, the reasoning, the model management and update and the learning. The decision state is composed of elements such as the selection of alternatives, its evaluation and optimization and the decision-making. The reconfiguration phase implements the decisions, selects the components and throws warnings and alarms. Finally, the communication phase is responsible for the command and data exchange between the several cognitive motors distributed along the network.

Inspired by the previous models, the cognitive cycle introduced in [13] is composed of the six phases shown in Fig. 5.
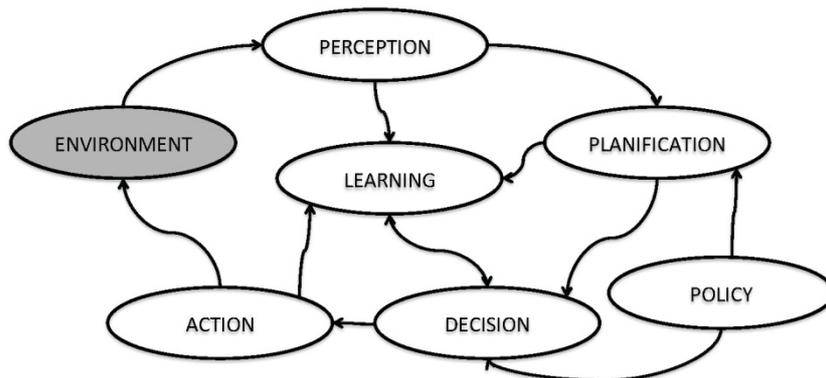

Fig. 5. Fortuna-Mohorcic reference cognitive cycle based on [13].

In this self-aware model, the nodes of the network have sensors to perceive the information of the environment. The captured observations are employed in the planning, but also to feed the learning module that builds and updates the model that remembers those observations. These models are used in the decision phase to choose the most adequate alternative according to the past experience. The planning module determines the potential strategies to follow based on the stored observations and policies. At last, the actions/reconfigurations related to the taken decision are accomplished. In this model, the learning module is connected to those others from which useful information can be extracted. This way, it can correlate and infer its own knowledge.

Besides the selected model, once the cognitive cycle and the techniques used in each phase are designed, it can be implemented in several ways. The cognitive capability may remain highly centralized or be absolutely distributed according to the design specifications. An example of cognitive cycle decentralization is the reconfigurable node model introduced in [14]. This model splits the phases into two entities (Fig. 6): the cognitive motor, which encompasses reasoning, learning and decision capabilities, and the reconfigurable node, which focuses on observation and action. The reconfigurable node is a structure composed of a set of hardware and software elements that are assembled by the cognitive motor according to the conclusions obtained from the information processing, and can be modified at convenience.
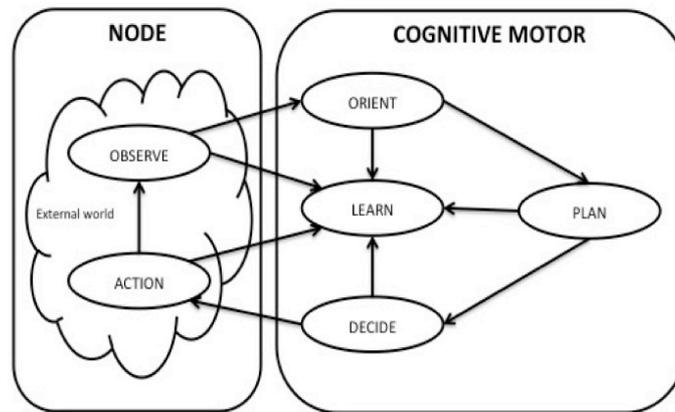

Fig. 6. Reconfigurable node architecture based on [14].

4 Application example of cognitive cycles to SESAME architecture

The SESAME EU-funded project contributes to the development of 5G focusing on three technological pillars: the use of NFV and MEC to move network intelligence

to the edge, (ii) the evolution of the Small Cell concept, and (iii) the promotion of multi-tenancy in communications infrastructures and services.

Based on these structural principles, SESAME suggests the architecture shown in Fig. 7. This architecture proposes the virtualization of small cell networks introducing the concept of Cloud Enabled Small Cell (CESC), a complete small cell that also contains a microserver offering, thus, computing, storage and radio resources. A number of CESCs form a cluster whose virtualized physical resources are shared and controlled by the VIM.
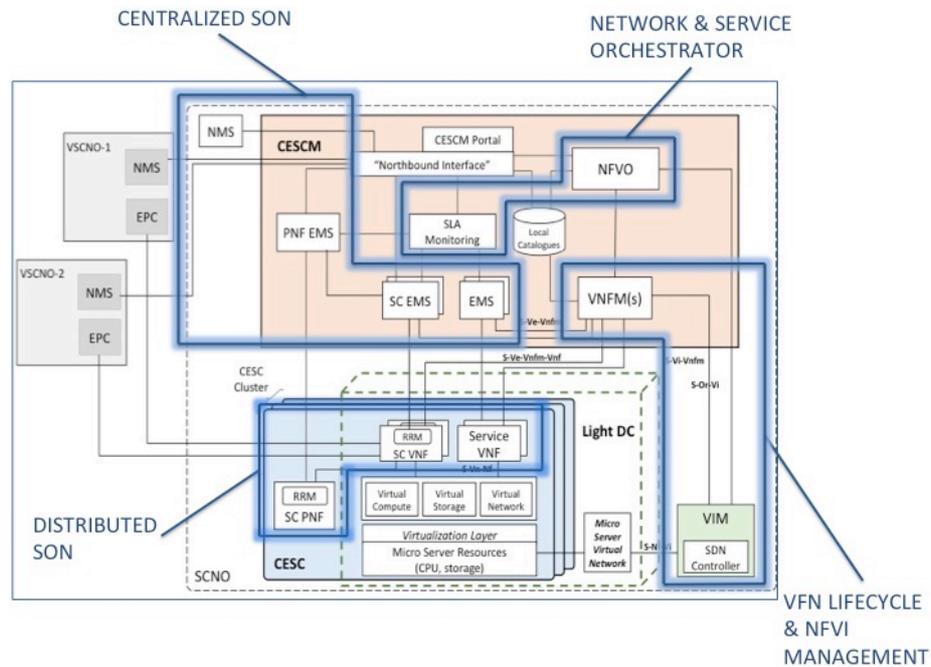

Fig. 7. SESAME overall architecture based on [1].

The CESC Manager (CESCM) is the central service management and orchestration component. A single instance of CESCM is able to operate over several CESC clusters, through the use of a dedicated VIM per cluster. Through virtualization, a Small Cell Network Operator (SCNO) can offer a cloud of resources that can be sliced to provide a logical instantiation of the network and, therefore, to enable multi-tenant services. The virtualized execution platform provided by the Light DC supports the VNF chains that implement the capabilities demanded to serve a requested network service (NS) by a tenant.

In summary, the Network Management System (NMS) arranges the logical network slices which enable SCNOs to provide Network-as-a-Service to the VSCNOs. Network slicing is supported by cloud edge-computing and NFV. The NFV Orchestrator (NFVO) provides management of those NFV services and is responsible for on-boarding of new network services and VNFs packages; NS lifecycle management;

global resource management; validation and authorization of network functions virtualization infrastructure resource requests [6]. Finally, the VIM is responsible for controlling and managing the compute, storage, and network resources to allocate the VNFs within the CESC cluster.

At this point, the application of Artificial Intelligence techniques makes sense in those elements of the architecture where decisions are taken. In the context of SESAME platform, the decision-makers should be placed in the control plane (VIM,VNFs and NFVO) and/or the NMS in the management plane, as highlighted in Fig. 7. In particular, the concept of cognitive cycle can be applied to implement learning processes fed by the data generated during system operation.

In principle, the philosophy of MEC combines better with a distributed cognitive model, like the aforementioned OOPDAL, CME or Fortuna-Mohorcic. Such a model observes and collects the local operation data generated in the edge and processes it heuristically to extract new knowledge in the form of enhanced local decision rules. This operation way reinforces the distinctive features of the MEC architecture, i.e. low latency, context awareness, minimized data transit and reduced network congestion.

On the other hand, a centralized cognitive model such as the reconfigurable node can gather all the information about the network operation in the cloud servers and take advantage of higher computational and storage resources. At this point, Big Data techniques can also be useful to process high volume of raw data, in order to achieve optimized decision rules that are later communicated to the edge.
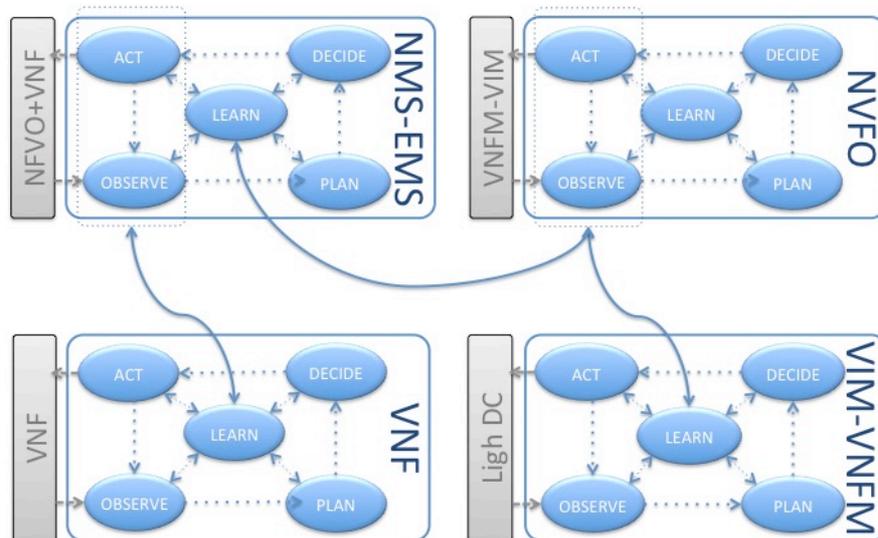
Fig. 8. Cloud-based multi-level cognitive cycle.

As a result of these considerations, we propose a combination of distributed and centralized model through a cloud-based multi-level cognitive cycle depicted in Fig. 8. The figure represents the four aforementioned decision points of the architecture.

At the lower level of this model, on the left, the Distributed SON (D-SON) module applies a simplified version of the OOPDAL cognitive module to manage the VNFs running in the CESC. The NS (VNFs) themselves observe their own environment and self-adapt their operation. On the right side of Fig. 8, the VIM applies the same cognitive cycle to support the arrangement of the physical resources of the Light DC. The indicators that feed the monitoring phase of the cognitive cycle are just those generated in each CESC cluster. In this context, the planning phase analyzes the possible alternatives to allocate and distribute the VNFs in the available resources to later apply a decision rule that will lead to an operation action. All these steps give feedback to the learning module that heuristically analyses the results and infers enhanced decision rules. This way, these actions are entirely performed in the CESC cluster, maintaining all the logic in the edge. In addition, the VNFM can request VNF scaling to the VIM according to pre-established rules, defining, thus, the Distributed NFV Management.

In the upper level, the decision points identified as Centralized SON and Network and Service Orchestration implement a version of the Reconfigurable Node Model. This version consists of the same simplified OOPDAL cycle acting as the cognitive motor and the reconfigurable node that includes in its observation phase the information from the lower levels that is uploaded through the cloud.

In the Network and Service Orchestration, the cognitive cycle supports the tasks of the NVFO aimed at composing service chains (constituted by two or more VNFs located either in one or more CESCs) and managing the deployment of VNFs over the Light DC. The NFVO monitors the performance level of the NSs for all the VSCNOs through the SLA Monitoring, and is able to change the capabilities of the network elements (eventually, VNFs). It must be remarked that the CESC Manager that allocates the NVFO can manage one or more CESC cluster, with the increment of sample data to process that it entails. In this context, this high volume of data may require the Big Data processing techniques that can be implemented in the big cloud servers.

Finally, this structure is repeated in the management plane, at the C-SON decision point. Here, the cognitive motor is implemented in the NMS-EMS, which are responsible for the management of the network slices that serve the NSs. Following the same rationale, the interaction with the lower level is bidirectional. The learning module of the cognitive cycle implemented in the NVFO acts as the reconfigurable node, providing an additional source of observations to feed the learning process and also receiving knowledge updates to enrich its operation with high level decision rules. Thus, the network slicing activity is refined complementing the OSS/BSS requirements with the NFV managing state for an optimized performance.

The proposed variation of the Reconfigurable Node Model supposes that the cognitive cycle at the lower levels can operate independently in a distributed way, but its operation can be uplifted when appropriate incorporating valuable knowledge from the upper levels. The consequence is a blurring-edge effect: the low-level distributed cognitive cycle reinforces the edge-computing concept maintaining the decision logic

close to the user, but at the same time, the reconfigurable node model adaptation moves the valuable operation information up to the cloud servers. This way, a bidirectional glocal effect is obtained. First, the cognitive cycle implemented within the distributed decision points can operate just with local information, but the generated knowledge feeds the cognitive motor at the centralized points leading to a global improvement over the performance of the network. In contrast, the knowledge results of the motor reconfigures as well the learning schema of the lower level, providing more accurate decision rules for future samples.

5   Conclusions

The application of cognitive models in the management of networked services has attracted significant effort for long time. This paradigm is especially relevant in mobile networks, where the dynamically variable context of use imposes more stringent requirements to the accurate management. Traditionally, the service and network management planes have been implemented in a rather isolated way, leading to non-optimal adaptation of services to network performance and of networks to service requirements. The evolution of different network softwarization techniques and its introduction into the design of future 5G mobile networks anticipates a revolution in the dynamic management of mobile services.

First, the SDN paradigm splits the control and data planes in data forwarding techniques. In this way, a centralized intelligent element is now capable of determining the most appropriate data paths for the different traffic flows. Additionally, the unification of different radio technologies in the framework of 5G networks introduces new capabilities and challenges in the cognitive management.

Second, the integration of service and network management endows the intelligent element with capabilities to perform consolidated cross-layer decisions. The advancement of mobile edge service technologies also brings new possibilities for an optimized coordination of the network and service data planes.

The use of cognitive cycles in this centralized unified service and network management system anticipates an enhanced performance of mobile services beyond the current capabilities of self organizing networks. However, the nature of end to end network services and the consideration of different network domains require also the inclusion of distributed cognitive cycles that locally work towards the overall enhancement.

The third technological building block analyzed in this paper is the inclusion of NFV techniques in the deployment of future 5G networks. The virtualization of network and service elements breaks the traditional limits of network-service performance maximization in each administrative domain, allowing the network provider to dynamically allocate the hardware resources where they are needed. Therefore, the paper analyses the possibility of building a higher-level cognitive cycle for the adaptive orchestration of virtualized resources, which in turn has an impact on the context information of the local cognitive management processes.

Different cognitive models have been presented and analyzed in the scope of 5G networks and specifically in the framework of dense cloud-enabled small cell deployments. As a result, the most appropriate cognitive schemes are discussed for the different decision making elements in the network-service provisioning.

Acknowledgements

References

1. H2020-SESAME: SESAME D2.2 v0.2: Overall System Architecture and Interfaces (2016)
2. ITU-T: FG IMT-2020: Report on Standards Gap Analysis (2015)
3. NGMN: 5G White paper (2015)
4. ETSI: Mobile-Edge Computing - Introductory Technical White Paper (2014)
5. Imran, A., Zoha, A.: Challenges in 5G: how to empower SON with big data for enabling 5G. IEEE Network 28(6) 27-33 (2014)
6. ETSI: ETSI GS NFV-MAN 001 v1.1.1: Network Functions Virtualisation (NFV); Management and Orchestration (2014)
7. Biglieri, E., Goldsmith, A.J., Greenstein, L.J., Mandayam, N.B., Poor, H.V.: Principles of Cognitive Radio (2012)
8. Clark, D.D., Partridge, C., Ramming, J.C., Wroclawski, J.T.: A knowledge plane for the internet. In Proc. 2003 conference on Applications, technologies, architectures, and protocols for computer communications - SIGCOMM '03, New York, New York, USA, ACM Press (2003)
9. Haigh, K.: AI technologies for tactical edge networks. IEEE Journal on Selected Areas in Communications keynote presentation for MobiHoc Workshop on Tactical Mobile Ad Hoc Networking. (2011)
10. Mitola, J.: Cognitive radio for flexible mobile multimedia communications. In Proc.1999 IEEE International Workshop on Mobile Multimedia Communications (Mo-MuC'99), pp. 3-10 (1999)
11. Strassner, J., Agoulmine, N., Lehtihet, E.: FOCALE: A Novel Autonomic Networking Architecture. In Proc. Latin American Autonomic Computing Symposium, Campo Grande, Brazil (2006)
12. Balamuralidhar, P., Prasad, R.: A Context Driven Architecture for Cognitive Radio Nodes. Wireless Personal Communications 45(3) 423-434 (2008)
13. Fortuna, C., Mohorcic, M.: Trends in the development of communication networks: Cognitive networks. Computer Networks 53(9) 1354-1376 (2009)
14. Sutton, P., Doyle, L.E., Nolan, K.E.: A Reconfigurable Platform for Cognitive Networks. In Proc. IEEE 2006 1st International Conference on Cognitive Radio Oriented Wireless Networks and Communications, pp. 1-5 (2006)