Scientific Research Publishing

Journal of Computer and Communications, 2023, 11, 33-62
https://www.scirp.org/journal/jcc
ISSN Online: 2327-5227
ISSN Print: 2327-5219

Artificial Intelligence Self-Organising (AI-SON) Frameworks for 5G-Enabled Networks: A Review

Delali Kwasi Dake

Department of ICT Education, University of Education, Winneba, Ghana
Email: dkdake@uew.edu.gh

How to cite this paper: Dake, D.K. (2023) Artificial Intelligence Self-Organising (AI-SON) Frameworks for 5G-Enabled Networks: A Review. *Journal of Computer and Communications*, **11**, 33-62.
https://doi.org/10.4236/jcc.2023.114003

Received: March 14, 2023
Accepted: April 22, 2023
Published: April 25, 2023

Abstract

The fifth generation (5G) networks will support the rapid emergence of Internet of Things (IoT) devices operating in a heterogeneous network (HetNet) system. These 5G-enabled IoT devices will result in a surge in data traffic for Mobile Network Operators (MNOs) to handle. At the same time, MNOs are preparing for a paradigm shift to decouple the control and forwarding plane in a Software-Defined Networking (SDN) architecture. Artificial Intelligence powered Self-Organising Networks (AI-SON) can fit into the SDN architecture by providing prediction and recommender systems to minimise costs in supporting the MNO's infrastructure. This paper presents a review report on AI-SON frameworks in 5G and SDN. The review considers the dynamic deployment and functions of the AI-SON frameworks, especially for SDN support and applications. Each module in the frameworks was discussed to ascertain its relevance based on the context of AI-SON and SDN integration. After examining each framework, the identified gaps are summarised as open issues for future works.

Keywords

Self-Organising Networks, Artificial Intelligence, Software-Defined Networks, 5G Networks, Big Data

1. Introduction

The fifth-generation (5G) wireless network is a significant evolution of 4G LTE networks. 5G is designed to support considerable growth in data and connectivity in a demanding networked society. There will be more than 50 billion linked devices by 2023, according to predictions [1], and 5G network traffic will amount to tens of Exabytes (1000^6 Bytes) each month. With this traffic, 5G networks are estimated to deliver 1000 times the capacity of current cellular systems [2]. 5G network realisation will support the massive increase in Machine to Machine (M2M) wireless communication systems [1] and the proliferation of applications requiring a lot of bandwidth, such as augmented reality, 3D videos, tactile internet, and virtual reality.

The Internet of Things (IoT) is a crucial component of the 5G infrastructure. IoT-based 5G networks are required to have high data rates, low latency, heterogeneous network existence and efficient spectrum management. To achieve these requirements, the integration of Artificial Intelligence (AI) is a necessity in the 5G ecosystem in analyzing the big data generated by IoT devices. This will help extract patterns and predict actions for the network and end devices to learn and improve from experience with minimal human intervention automatically.

This paper presents a detailed review of suggested traffic intelligence gathering interventions in 5G networks. The article further details the categories of the study under Artificial Intelligence and Self-Organising Networks (AI-SON) and AI-SON integrated into Software-Defined Networks (SDN). This review comprehensively explains the merits of Artificial Intelligence and SDN models in 5G and suggests new standards for efficient intelligence gathering. All frameworks reviewed were redrawn for clarity.

2. Artificial Intelligence and SON (AI-SON)

The dynamic resource allocation in 5G networks demands AI-integrated automation of network management processes. In this regard, Self-Organising Networks (SON) have evolved from traditional manual management processes to cater for the automation gap [3]. The SON is designed for the Radio Access Network (RAN) to have self-planning, configuration, management, optimisation and healing capabilities [4]. Mobile Network Operators (MNOs) will need to deploy wireless communications systems that can handle 1000 times the current traffic volume, one trillion linked devices, a wide variety of use cases, and enhanced specific performance needs in the future [5] [6].

The reviewed literature under this section includes SON models for the effective integration of AI for Big Data analytics to deal with the increasing traffic expected from the 5G-RAN deployment by the MNO.

2.1. AI-Based Framework for 5G Network Planning and Operation

Pérez-Romero *et al.* [7] developed a 5G network design and operation infrastructure based on AI to support the concept of Self-Organised Networks (SONs) in perceiving and analyzing the 5G ecosystem intelligently. As illustrated in **Figure 1**, the framework acts on input data from its environment and processes it for appropriate network actions to aid the MNO's decisions validated through decision support systems.

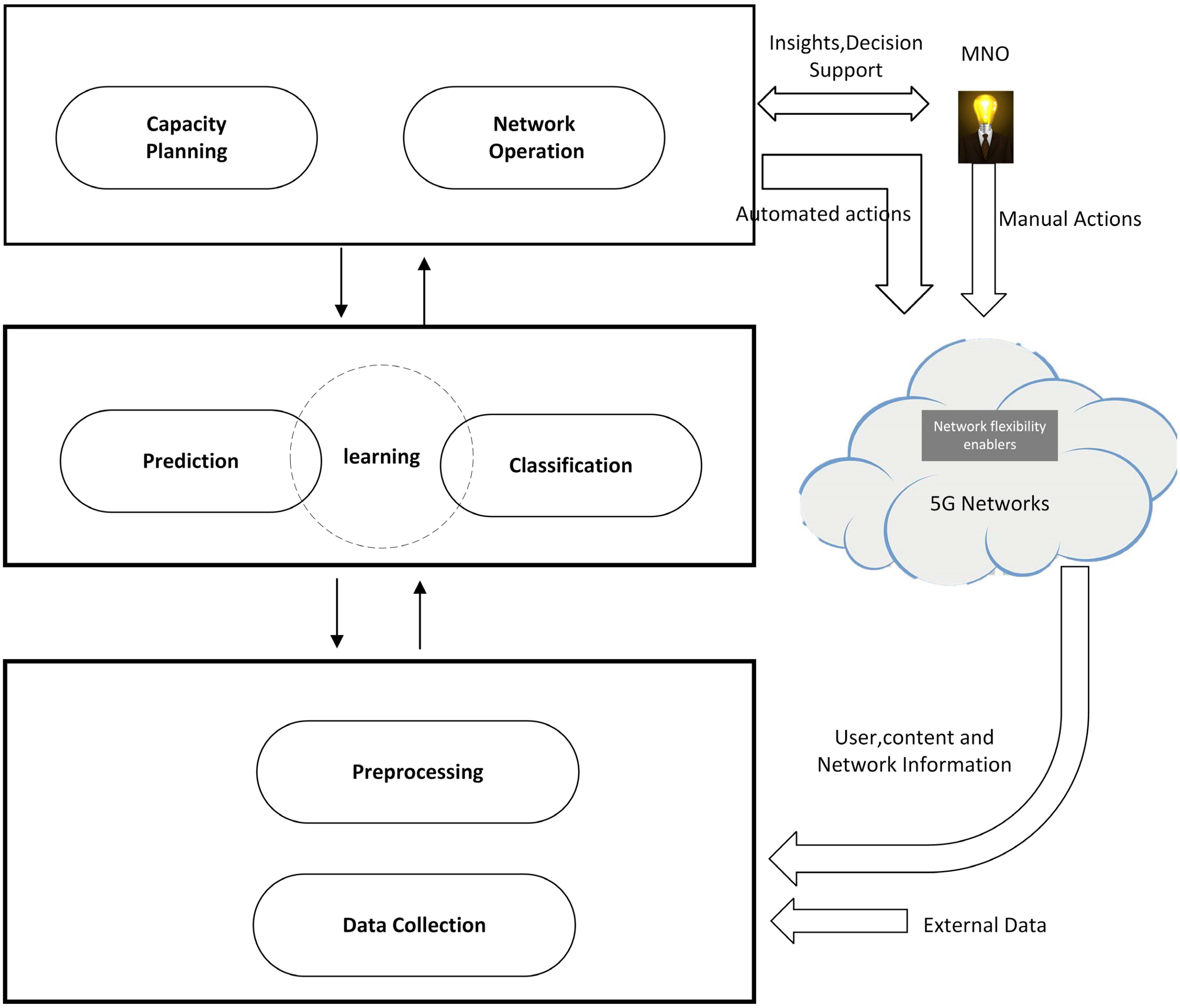

Figure 1. AI-based framework for 5G network planning and operation [7].

Data acquisition and pre-processing is the first step in the Pérez-Romero *et al.* [7] paradigm. MNOs gather complex data from several systems for managing customers, networks, billing, inventory, server management, devices for deep packet inspection, and databases tailored to specific applications [8]. The data gathered are categorised into:

- Network data defines network behaviour comprising radio-related measures taken by base stations and terminals, consumption data at network setups, network routes and nodes, network performance indicators, and QoS measurements.
- User data comprises subscriber profiles accessing the network, demographics, pricing and strategies, devices, capacities for subscription, and used applications.
- Content data, including information linked to the network's applications.
- External data includes data derived from non-MNO sources.

These data sources are heterogeneous and pre-processed to eliminate noisy and inconsistent data. The data cleaning process includes data integration, selection and transformation for specific mining.

As part of the Pérez-Romero *et al.* [7] paradigm, knowledge discovery is used in the second stage to infer conclusions from the pre-processed data and create models that reflect the pertinent knowledge to guide operational and planning decisions.

The third stage involves knowledge exploitation. This includes a set of recommendations to be used with the network and the retrieved information for prediction. The framework only allows processes for capacity planning and network operation. The network operation processes include operations performed on deployed resources to regulate and optimise network behaviour. Capacity planning, on the other hand, entails providing the necessary network resources to handle traffic demands.

The last stage involves the network flexibility enablers. SDN and Network Functions Virtualisation (NFV) technologies facilitate the AI-based framework. The SDN controller will aid the implementation of insight decisions for the MNOs to enforce automatic or manual actions. Through NFV, several tenants can share network infrastructure, allowing for flexibility in changing network configuration and architecture.

2.2. A Data-Driven Framework for Personalised QoE

A data-dependent framework for improved personalised QoE for 5G networks was developed by Wang *et al.* [9] with a two-step QoE modelling technique to capture the link between end users and services and the user's subjectivity toward a particular service. The proposed architecture:

- Contains a monitor to capture real-time data about user programs and the QoS status
- Has a data mining engine that forecasts user expectations regarding the application in use
- Maintains a suitable QoE by managing the transmission resources based on the QoS state and the projected demand.

As illustrated in **Figure 2**, the architecture has two parts; the offline and online parts. The offline component is in charge of training and transforming the user profile prediction model. The model is per-user per-service. One component of this model is the Subjective Data Collector (SDC). As depicted in **Figure 3**, the fundamental component of the data collection component is a mobile agent placed on end-user devices for QoS monitoring, contextual monitoring, and experience monitoring.

- The QoS monitoring entity collects data on various technical parameters. These parameters include information about the device (such as screen size and operating system), the network (including access type, jitter, throughput, and delay), and the application (such as application type).
- The contextual monitoring entity is in charge of gathering the application's

context information, which includes the device's location, mobility, and sensor data.

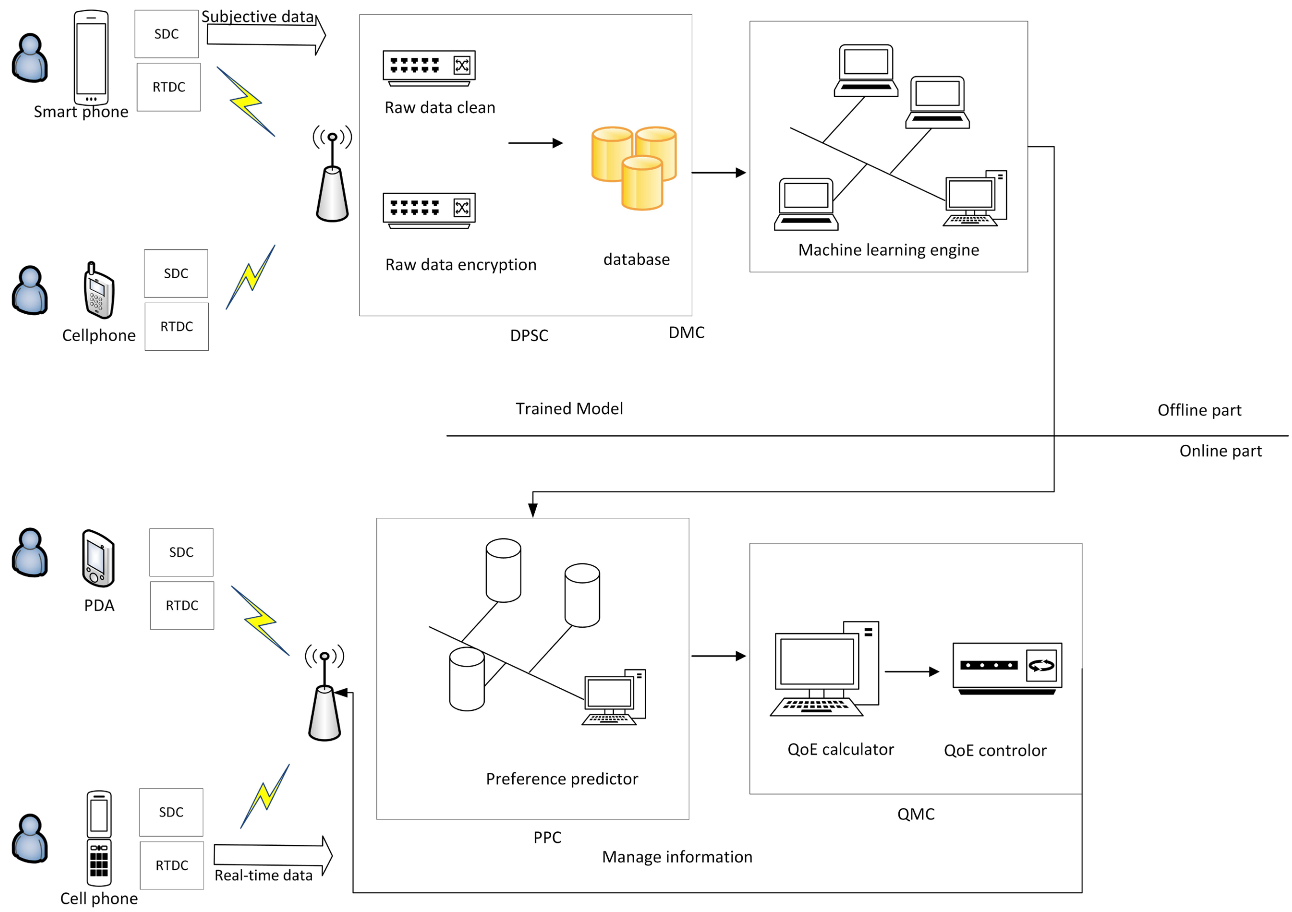

Figure 2. A data-driven architecture for personalised QoE [9].

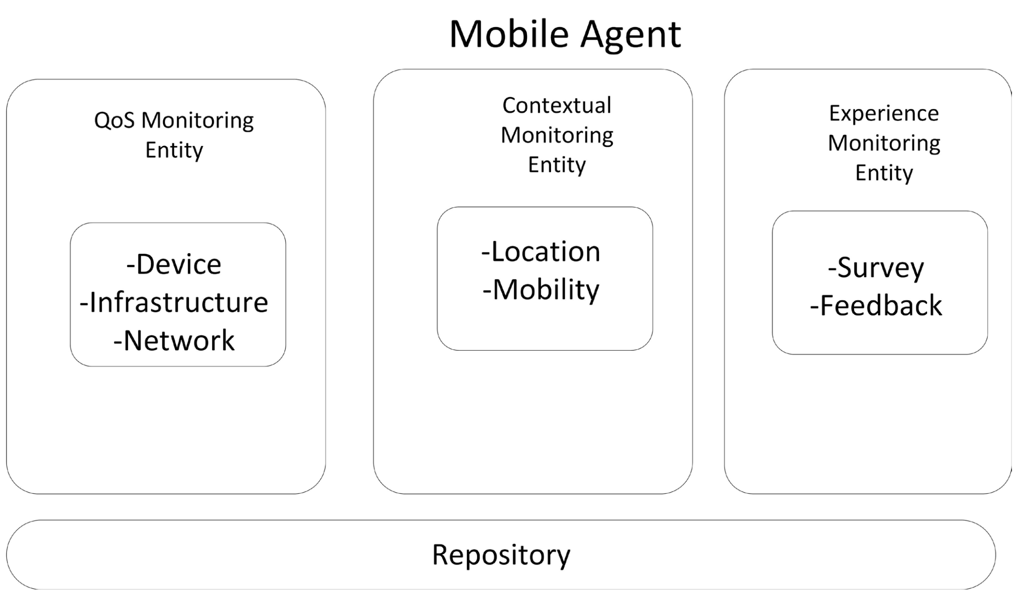

Figure 3. The structure of the mobile agent [9].

- The experience monitoring entity communicates with users by collecting explicit feedback through questionnaires.

The data from the SDC is synchronised by the Data Processing and Storing Component (DPSC). The DPSC's primary duty is to pre-process data before storing it in databases for mining. The model is trained using the Data Mining Component (DMC) to understand user preferences. This forms the system's central component. One or more data mining models are constructed and trained using the DPSC data. Subsequently, the online component will make use of the trained models.

Wang *et al.* [9]'s architecture's online component gathers real-time data about users and related services. Predictions are made offline using data mining algorithms that have been trained. The online part has three components:

- Real-Time Data Collector (RTDC). The RTDC captures current user information, the user's services and the network resources.
- Preference Prediction Component (PPC). A preference predictor is the PPC's central component. The predictor's job is to process data from PPC using the appropriate data mining model trained offline to forecast a specific user preference.
- QoE Management Component (QMC): The QMC receives user preferences derived from the PPC. The QMC has a QoE controller and calculator. The controller is aware of the objective network's state and system users' arbitrary preferences. It is possible to carry out QoE optimisation, such as maximising overall QoE or enhancing QoE equity for all users. With this design, specific QoE user optimisation at QMC is also possible.

2.3. Knowledge Acquisition Framework for 5G Environment

Monge *et al.* [10] proposed an automated analytic approach for inferring knowledge in 5G networks to determine network conditions and evaluate potential circumstances that could interrupt network operations. The framework uses automatic metrics discovery, pattern recognition, and prediction approaches designed with the Endsley situational awareness paradigm. As depicted in **Figure 4**, the intelligent framework consists of successively pipelined architectural elements: Discovery, Onboarding of use cases, Notification, Pattern Recognition, Predictions, Adaptive Thresholding, and Knowledge Inference.

The discovery component acquires data, represented as facts, collected by network sensors in a 5G architecture through surveillance, data collection, and connected operations. These monitored metrics are facts obtained as Aggregated Data Bundles (ADBs), summarising acquired observations. The Inference Engine acquires new knowledge based on the collected facts stored in the working memory, enabling the inference of network status conclusions by applying rules set in the Knowledge Base. The Knowledge Base is populated with information obtained from use case definitions and acts as procedural knowledge and system inference. For inferring new knowledge, the Prediction, Pattern Recognition,

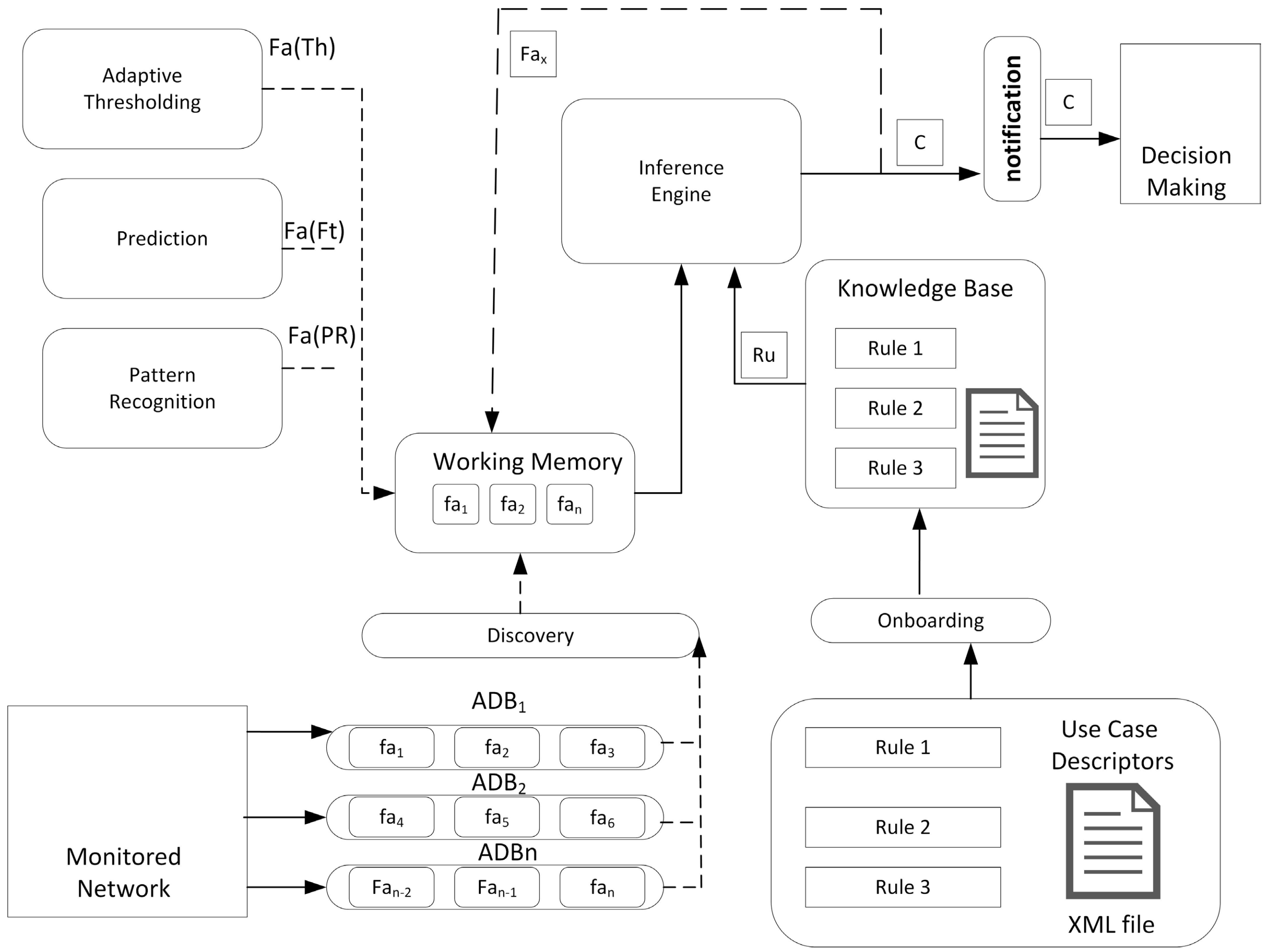

Figure 4. Knowledge acquisition framework for 5G environment [10].

and Adaptive Thresholding components generate new factual knowledge internally. The Prediction and Adaptive Thresholding components of the Monge *et al.* [10] framework permit a situational awareness projection of the network by computing predictive metrics and forecasting intervals to enable pro-action reactions over the projected scenarios.

The component for pattern recognition incorporates several AI concepts, including machine learning, data mining, and classification.

2.4. Supervised and Unsupervised Learning Framework

Fu *et al.* [11] proposed an AI framework for 5G wireless networks to manage network traffic. The paradigm for traffic engineering in the 5G network is based on supervised learning and unsupervised learning perspectives.

As illustrated in **Figure 5**, the supervised learning framework trains a model using training samples labelled by a knowledge base. In classification and regression scenarios, training data form a mapping between the inputs and desired output labels. The trained model then infers reasonably for new inputs. The unsupervised learning perspective identifies and comprehends the data structure

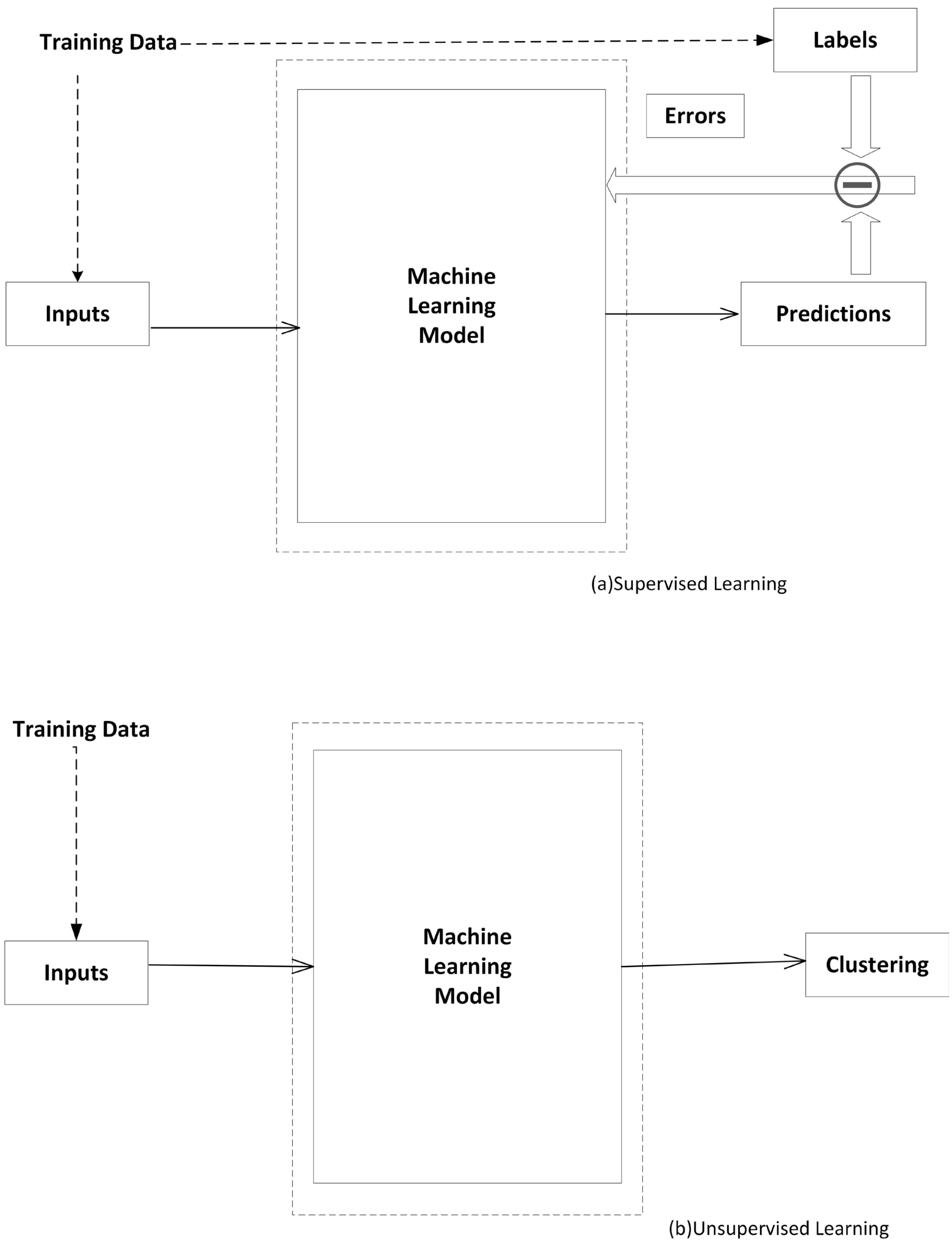

Figure 5. Supervised and unsupervised learning framework [11].

for a clustering problem to avoid the difficulty in learning from data labels. The unsupervised learning for 5G will facilitate data distribution modelling across the 5G networks, congestion and traffic conditions scenarios. This helps in scheduling and configuration for adaptive network traffic and topology changes in 5G.

2.5. A Machine Learning Framework for Resource Allocation

A machine-learning approach for resource allocation using cloud computing was proposed by Wang *et al.* [12]. As shown in **Figure 6**, the framework has a

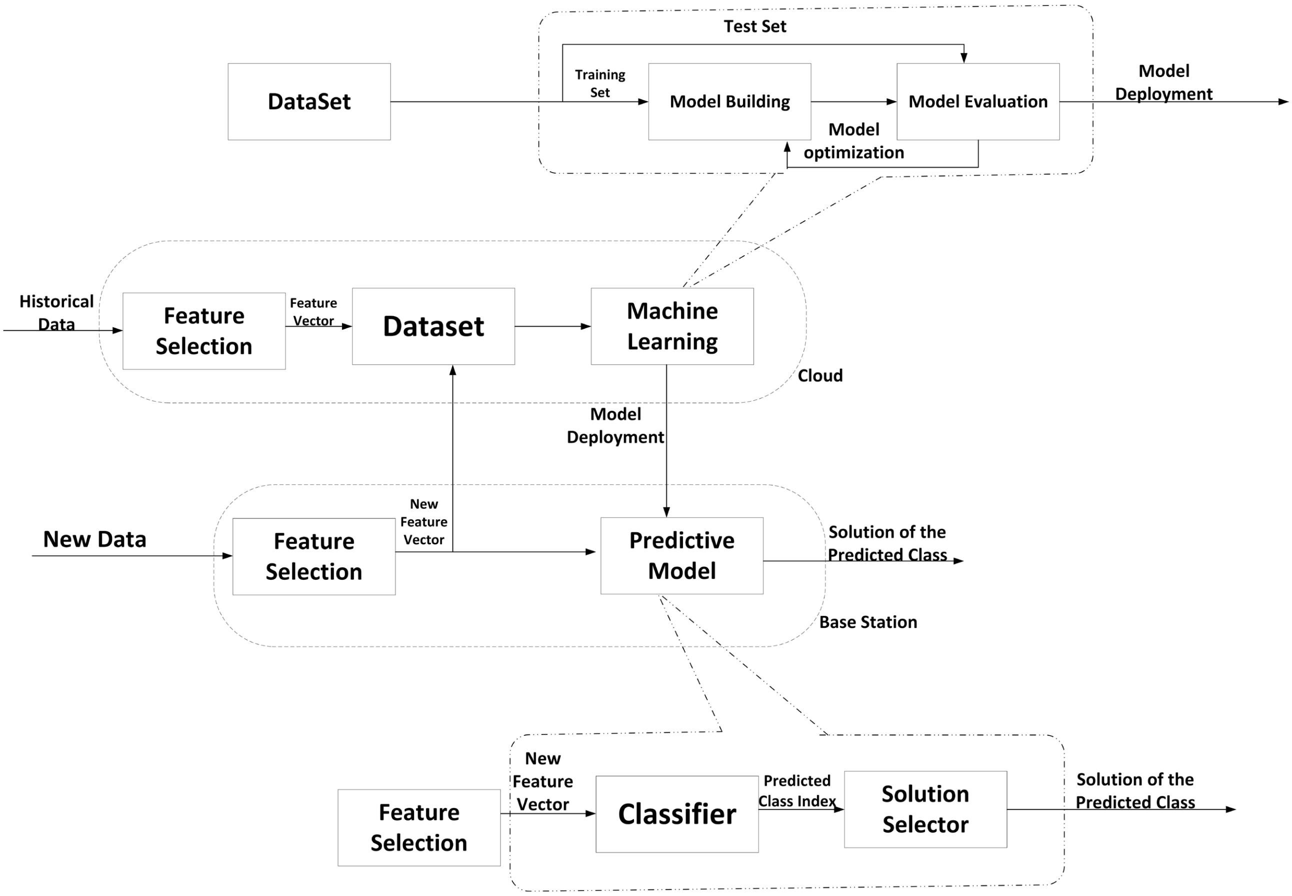

Figure 6. A machine learning framework for resource allocation [12].

cloud component for storing historical data based on scenarios. The historical data collected has a lot of attributes, and filtering is relevant to specific resource allocation.

Since learning from a massive amount of raw data with several attributes requires a lot of memory and processing resources, feature selection is essential to exclude irrelevant attributes. This forms part of the pre-processing of this framework, with an estimated 70% - 90% of training set data. A supervised learning algorithm is applied on the training set to uncover hidden similarities in historical data. Then, a predictive model is built to determine resource allocation for future unforeseen events. One class is created from all training feature vectors with the same solution, and each class has its inner solution. The resource allocation framework is changed into a multi-class classification problem. Two procedures are used to construct a predictive model. The first is to forecast the class for hypothetical future situations, and the second is to analyze test set data to assess the prediction model.

Using the backhaul lines, the Base Station (BS) receives the constructed predictive model and the related solutions for all classes. At the BS, new feature vectors formed will serve as inputs into the predictive model used to allocate the radio resources. The feature vector will be gathered and temporarily stored at BS before being forwarded to the cloud to update the dataset

2.6. BDD Network Optimisation Framework

As illustrated in Figure 7, Zheng *et al.* [13] proposed a Big Data-Driven (BDD) paradigm for network optimisation in 5G that encompasses Big Data gathering, storage management, data analytics, and network optimisation. Big Data sources from the framework include User Equipment (UEs), Radio Access Networks (RAN), Core Networks (CN), and Internet Service Providers (ISPs). MNOs under the proposed framework require a scalable huge storage infrastructure to administer acquired multi-source, heterogeneous, real-time and massive data.

Zheng *et al.* [13] BDD framework has a Big Data analytics module that monitors and analyses real-time data across users, mobile networks and service providers. Resource allocation can be used in conjunction with MNO Analytics deployment to forecast where and how users will utilise the mobile network, which could lead to potential traffic congestion at specific places. The operators can devote extra resources to addressing the peak traffic and preserving user QoE with the help of anticipated information from data analytics.

2.7. SELFNET

SELFNET [14] uses actuators to prevent potential issues and sensors to keep track of specific network information.

As depicted in Figure 8, the SON Autonomic Layer provides network intelligence. It consists of two sublayers: Monitor and Analyser Sublayer and Autonomic

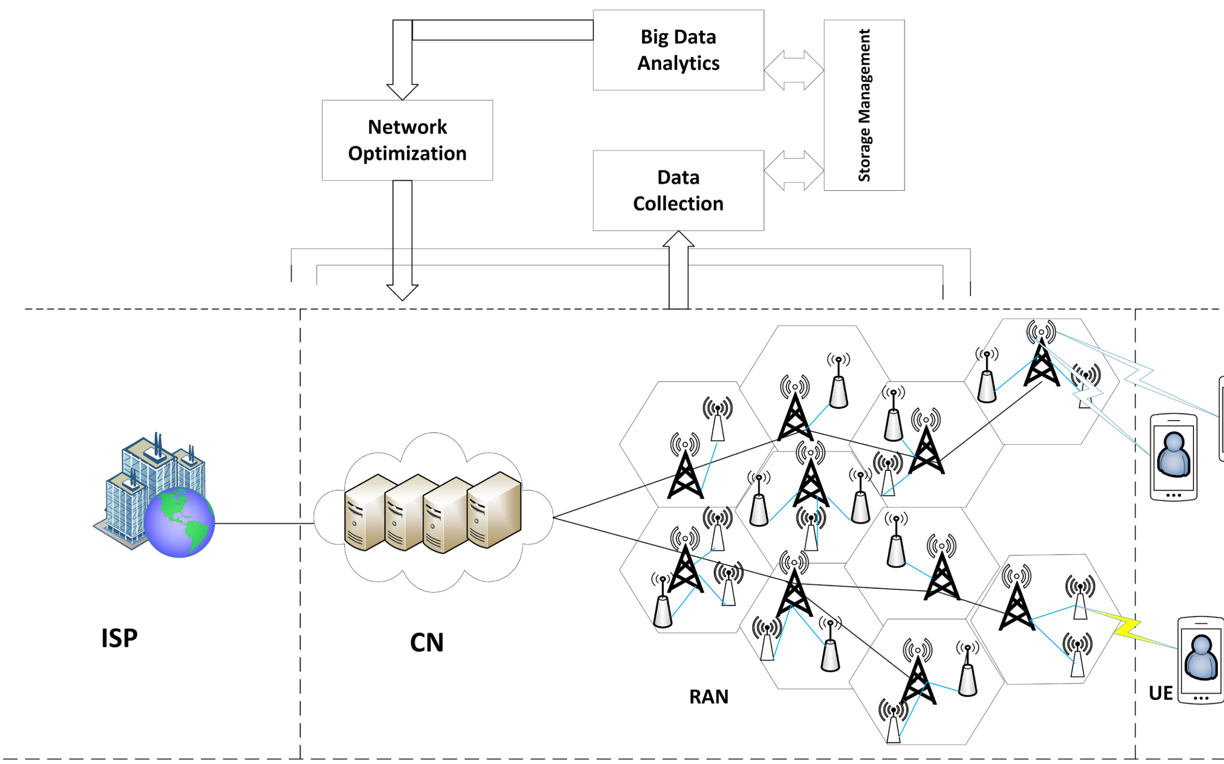

Figure 7. Proposed BDD network optimisation framework [13].

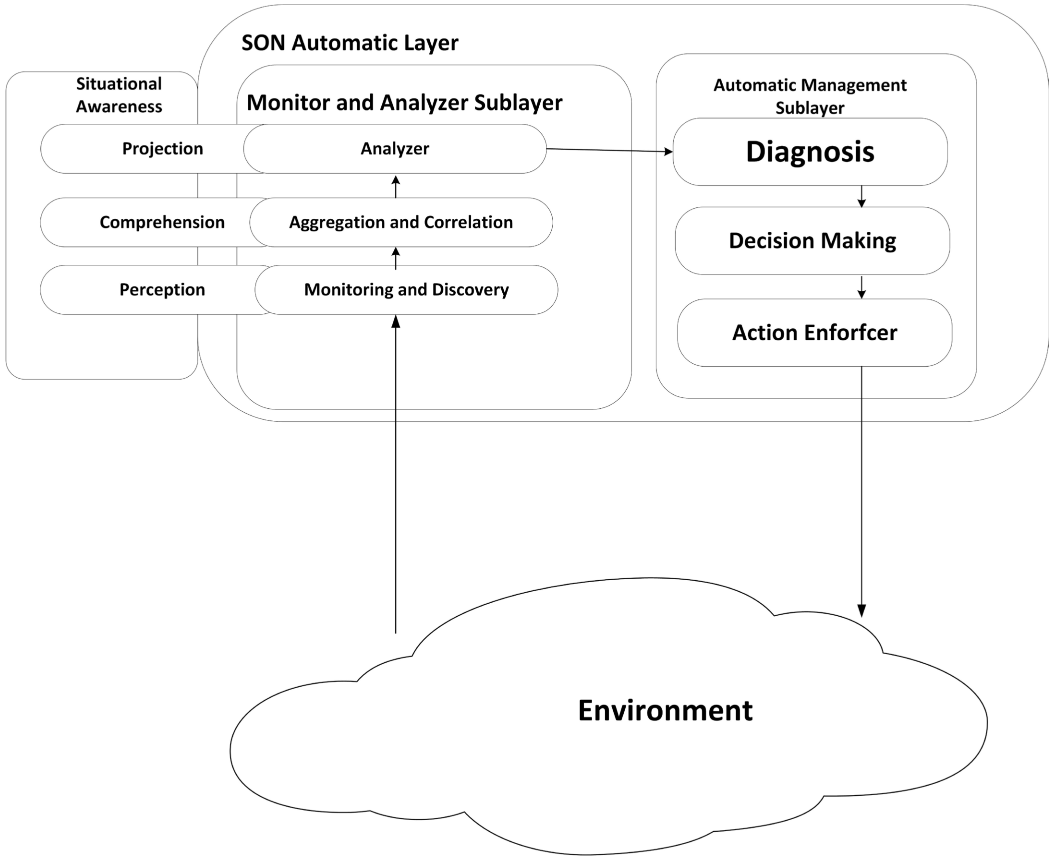

Figure 8. Endsley vs. SELFNET autonomic layer [14].

Management Sublayer. The Monitor and Analyser Sublayer, which follows the Endsley Situational Awareness Principles, comprises Monitoring and Discovery, Aggregation and Correlation and Analyser modules. These modules are linked to the Perception, Comprehension and Projection functions of Endsley principles. The Analyser module infers data from monitored metrics based on the knowledge provided by the Monitoring and Aggregation phases of the SON Autonomous Layer. After intelligence is received from the Analyser module, the Autonomic Management Sublayer of the SON Autonomic Layer activates Diagnosis, Decision Making, and Action Enforcement policies.

The brain of SELFNET, as depicted in **Figure 9**, is the Analyser module architecture, which is responsible for information collecting and permits proactive network infrastructure responses. The SELFNET Analyser consists of eight fundamental components. Sets 4, 5, 6, and 8 are associated with reasoning, 1 and 3 with projection, and 7 with the administration of use cases.

3. AI-SON Integrated in Software-Defined Network (SDN)

SDN is an essential and growing network architecture that decouples network control from data forwarding by direct programming [15]. By decoupling the control plane from the data plane, SDN provides better network programming control [16] [17]. There are three distinct levels:

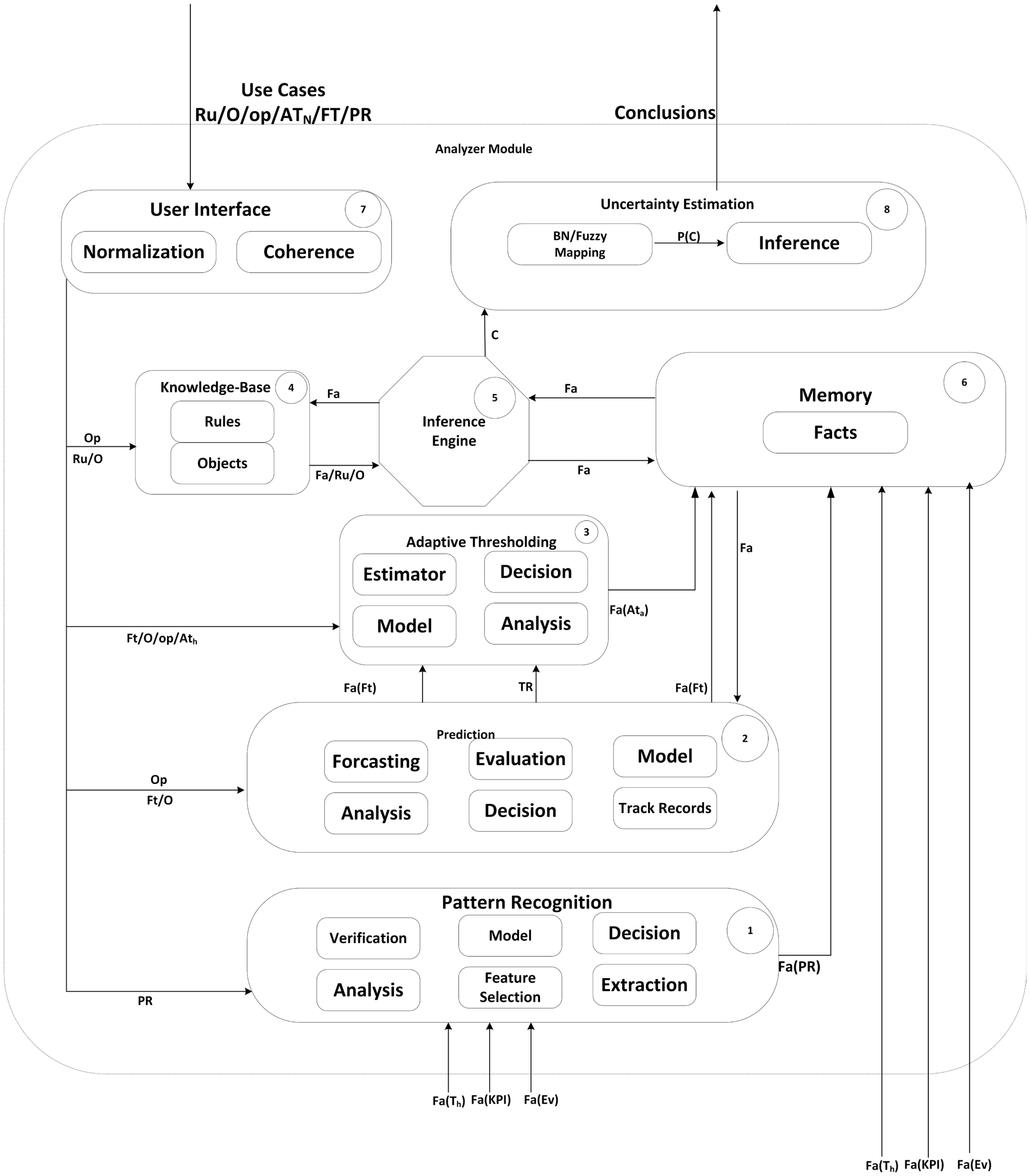

Figure 9. Analyzer module architecture [14].

- Application Layer: Software applications reside in this layer and communicate with the control layer
- Control Layer [18]: As the SDN's central hub, the controller receives requests from the application layer, maintains a logical perspective of the entire network, and manages network devices using industry-standard protocols.

- Data-plane Layer: These infrastructure components can be programmed with standardised interfaces in an SDN.

The researchers in this section proposed the possible integration of AI-SON as a component in SDN architecture at the Application, Control or Data-plane Layer.

3.1. SDN-Based Intelligent Model for HetNet

Sun *et al.* [19] proposed an SDN-based smart framework for efficiently managing the heterogeneous network (HetNet) infrastructure and resources. In this paradigm, the control plane is independent of the data plane [20], with SDN recognising the network as an operating system to abstract applications from the hardware. As depicted in **Figure 10**, Sun *et al.* [19] framework has one feature with the same infrastructure as 4G LTE, consisting of a core network (CN) and a RAN. The system's CN has three components: the Serving Gateway (S-GW), the Mobility Management Entity (MME), and the Packet Data Network (PDN) Gateway (P-GW).

Pico, Femto, and macrocell BSs form the evolved-universal terrestrial radio access network (E-UTRAN), a type of RAN. The framework incorporates SDN with a high-level abstract to which underlying network resources are mapped automatically with 5G core networks. With Sun *et al.* [19] architecture, the E-UTRAN and the eNBs are virtually implemented and managed centrally by

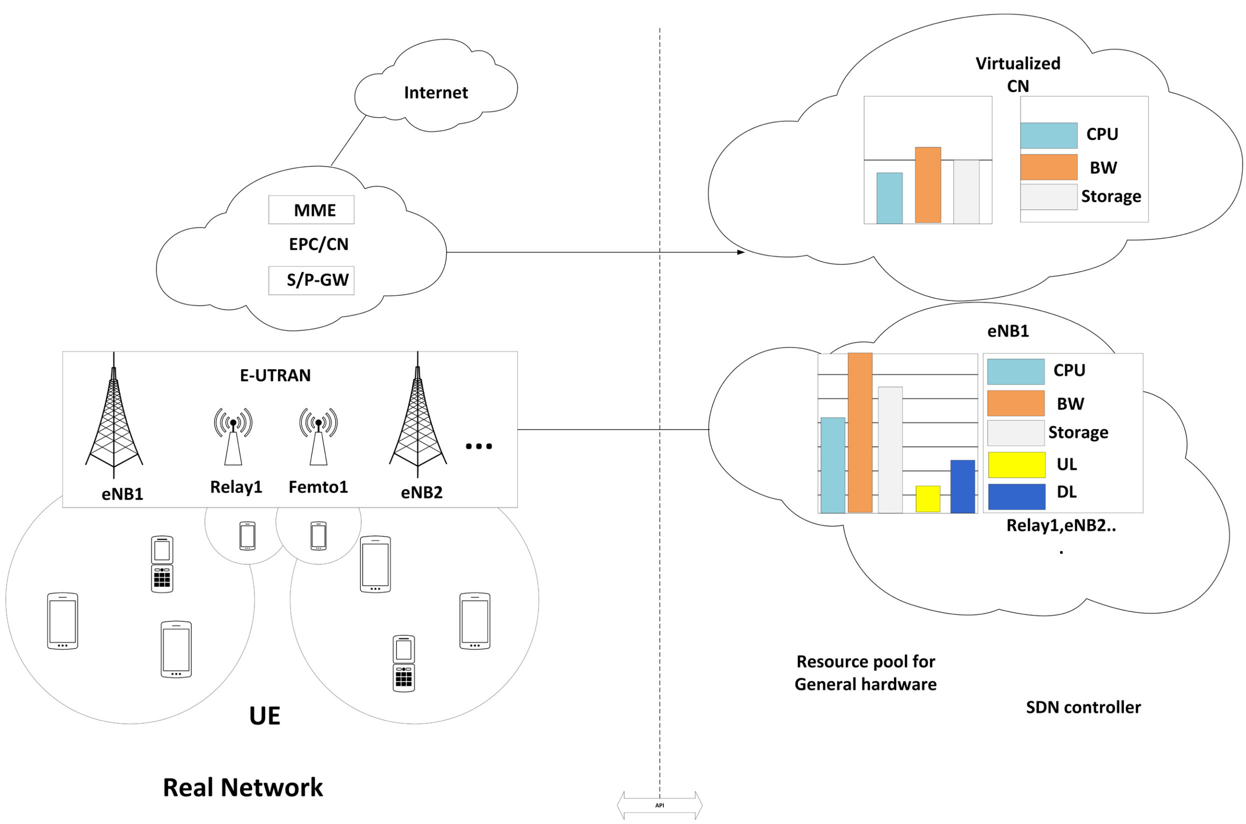

Figure 10. Proposed system model of intelligent SDN [20].

SDN controllers. The SDN controller is deployed physically on centralised servers to abstract present resource utilisation and intelligently operate network parts via APIs.

3.2. Self-Healing Framework for SDN

A Self-Healing design, depicted in Figure 11, was presented by Sánchez *et al.* [21] to guarantee the continuous operation of all service nodes. Using network observations to execute recovery activities, the Self-Healing framework affects the three SDN planes and the service plane.

SDN [21]

The proposed framework interacts with the SDN architecture by:

- End-to-end service reconfiguration
- Dynamically orchestrating SDN applications in response to faults or changing conditions
- Using the SDN controller to reprogramme the data plane
- Acting directly on the data plane to set configurations while performing manual installation that the controller cannot accomplish independently.

When a service fails, the application, control, and data plane symptoms are retrieved to look for correlations.

3.3. Cellular SDN (CSDN)

Bradai *et al.* [22] proposed a Cellular SDN (CSDN) architecture that uses SDN and NFV for dynamic resource orchestration as illustrated in Figure 12. This architecture streamlines network management and control by creating flexible,

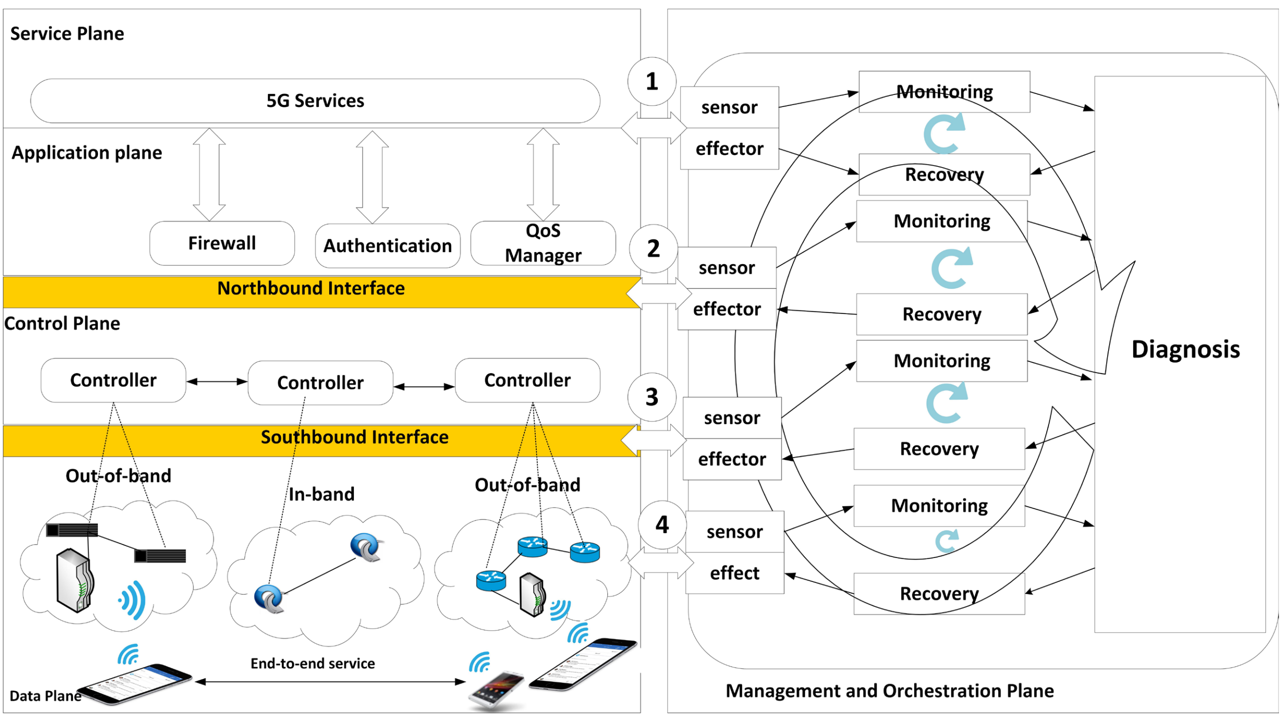

Figure 11. Self-Healing framework for SDN [21].

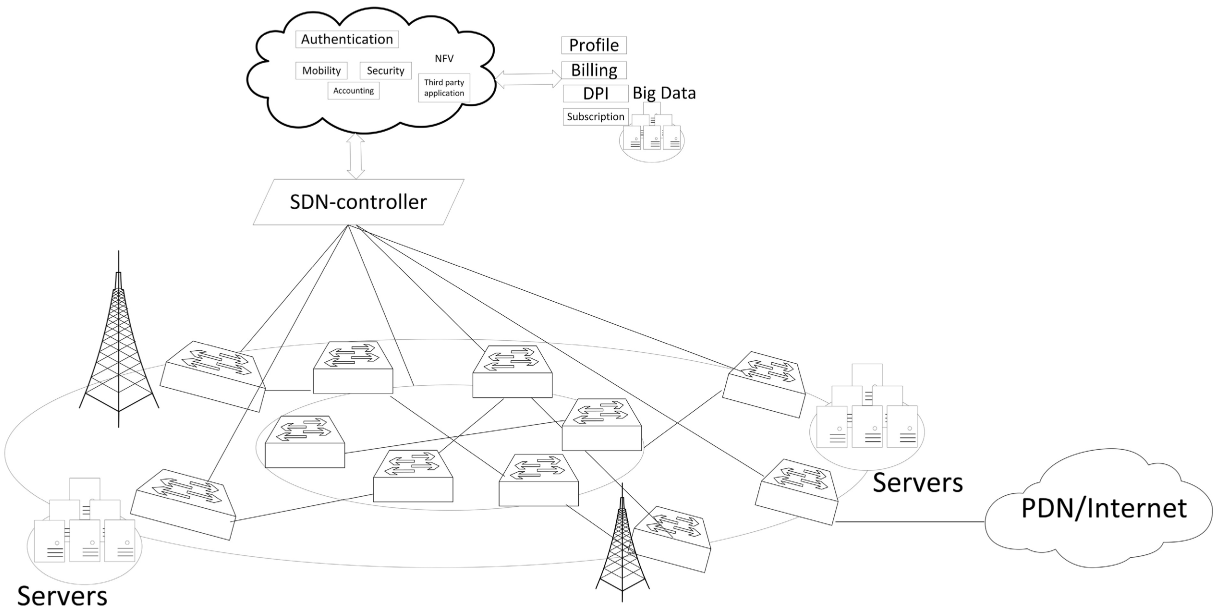

Figure 12. CSDN architecture [22].

open, and programmable new services. The CSDN design incorporates intelligent services, enabling Mobile Service Providers (MSPs) to implement subscriber policy and profile-aware service provisioning.

The MSP relies on data analytics to make decisions to realize intelligent services. A context data repository (CDR) includes network data, user profiles, and usage data. In addition, network data such as traffic load, bandwidth availability, wireless channel information, and network health data are collected. Together with the network data, these constitute the user-centric and network-centric data necessary for intelligent resource allocation and provisioning.

As depicted in **Figure 13**, the CSDN model is augmented with an extra knowledge layer that enables the MSP to obtain insight into the network's intelligent vision and user actions. As illustrated in **Figure 14**, the Knowledge plane consists of three functional blocks and two interfaces. The data acquisition block enables the collection of data from either the network or an application running in the CSDN application layer.

The next step involves data analysis for valuable insights to the network operator for comprehensive decision. The data organisation and management block follow this to facilitate knowledge exploitation.

3.4. Intelligent IoT-Based 5G Ecosystem

Javaid *et al.* [23] presented a firmware-based approach for integrating intelligence into IoT devices. The IoT-based 5G ecosystem depicted in **Figure 15** enables the application of AI techniques to Big Data for prediction and real-time analysis. They proposed the integration of intelligence at the application level for proactive and real-time decisions during runtime. Introducing AI into the firmware's essential components provides a secure environment for the execution of

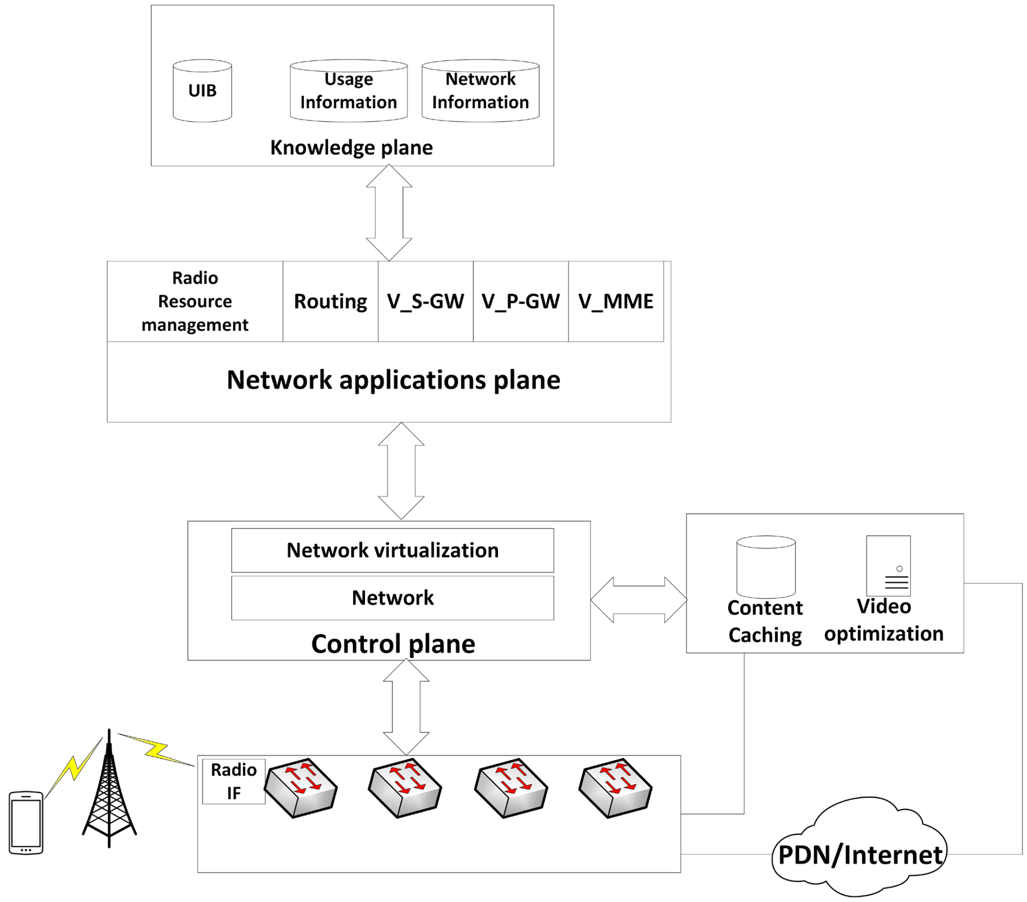

Figure 13. Detailed CSDN architecture [22].

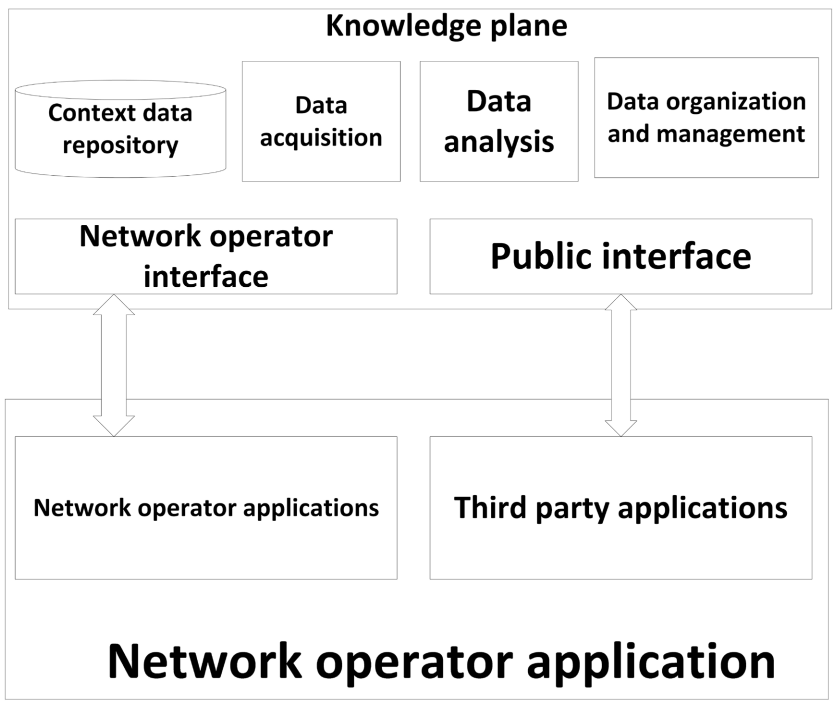

Figure 14. Knowledge plane of CSDN [22].

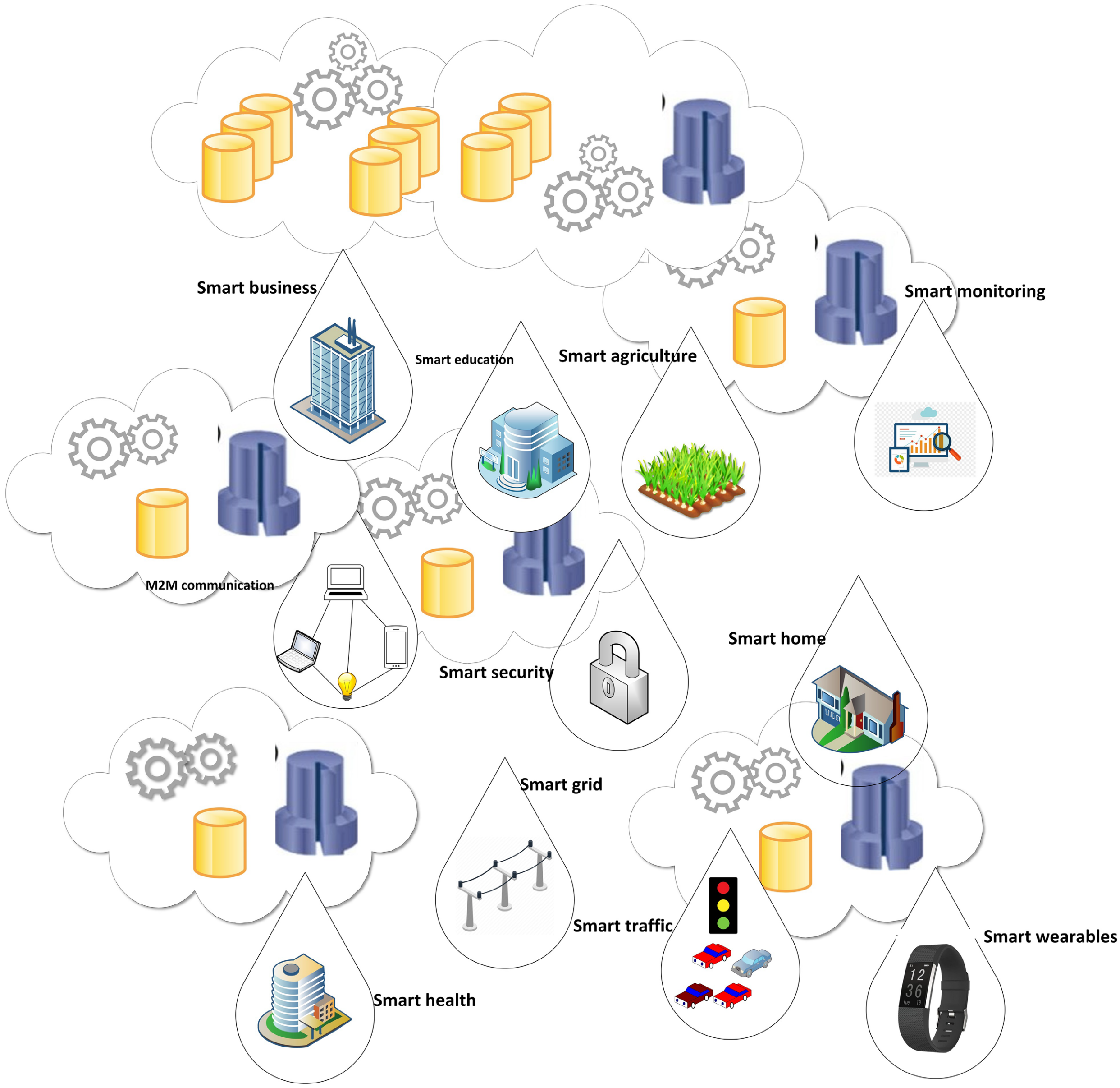

Figure 15. IoT-based 5G ecosystem [23].

applications to make informed judgments. As illustrated in **Figure 16**, AI is incorporated at the firmware level for data structuring and API runtime management with a keen understanding.

This design integrates 5G ecosystem intelligence at the application platform as a service (aPaaS), which encompasses operating system and middleware, communication patterns, and network infrastructure coexistence. The intelligent IoT with firmware can communicate via IoT switches and the cloud via an IoT gateway.

3.5. Future Intelligent Network Framework

Xu *et al.* [24] proposed a Future Intelligent Network (FINE) framework based on

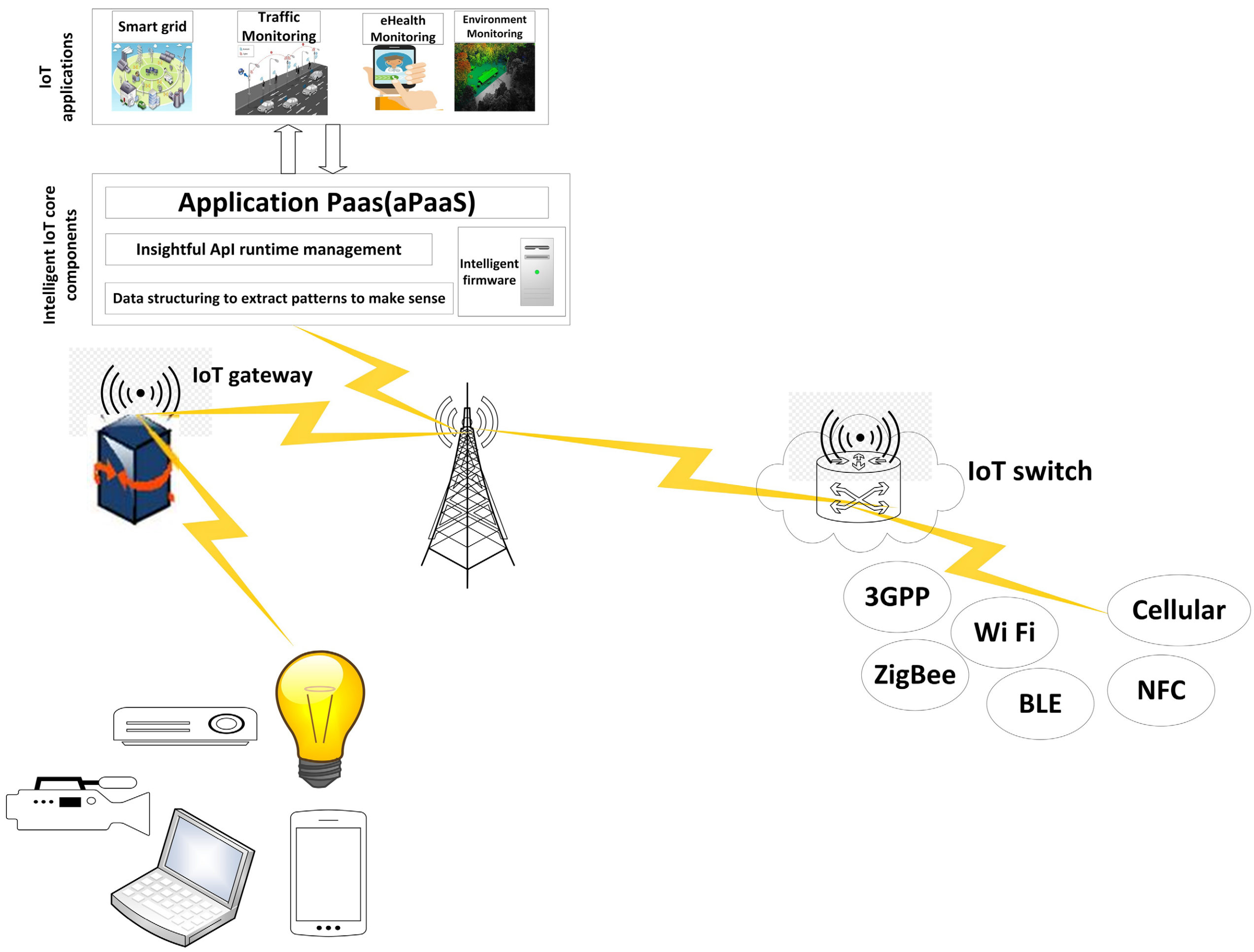

Figure 16. Intelligent IoT-based 5G ecosystem [23].

SDN and NFV technologies. The framework relies on Deep Packet Inspection (DPI) Systems [25] for data collection, including the running state of network equipment, resource usage and the quality of services. This serves as Big Data for the FINE framework and is shown in Figure 17. The DPI is installed on each network component, and the information collected by the DPIs is transmitted to the Big Data module of the FINE framework.

As depicted in Figure 18, the system architecture of the FINE framework consists of three planes: the Intelligence Plane, the Agent Plane, and the Business Plane. The FINE framework's basic layer of the Intelligence Plane gets Big Data from DPI systems.

As shown in Figure 19, the Intelligence Plane is responsible for providing intelligence for the entire FINE. This plane is composed of the basic layer, the core layer, the platform layer, the application and terminal layer and the solution layer. The basic layer receives the big data and the network status data of all the equipment, applications and services by relying on the network base and computing base. The core layer in FINE provides intelligent algorithms in the intelligent plane. This forms the kernel of the FINE framework.

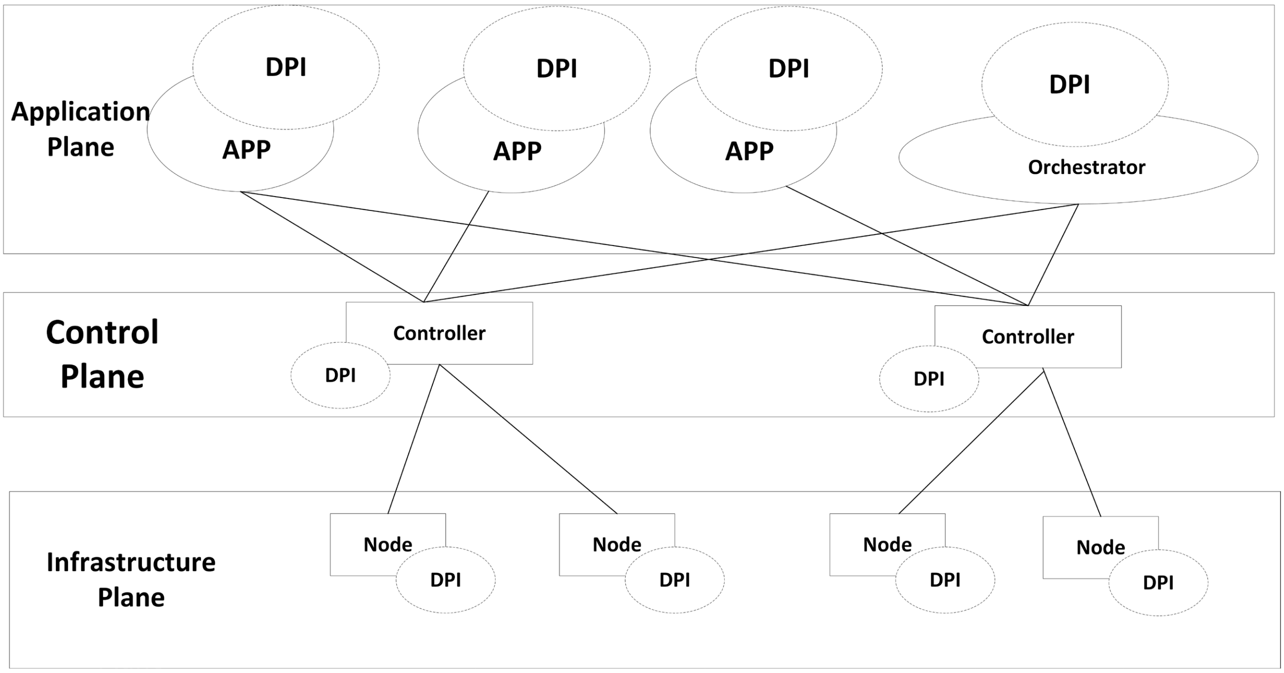

Figure 17. An SDN/NFV with DPI [24].

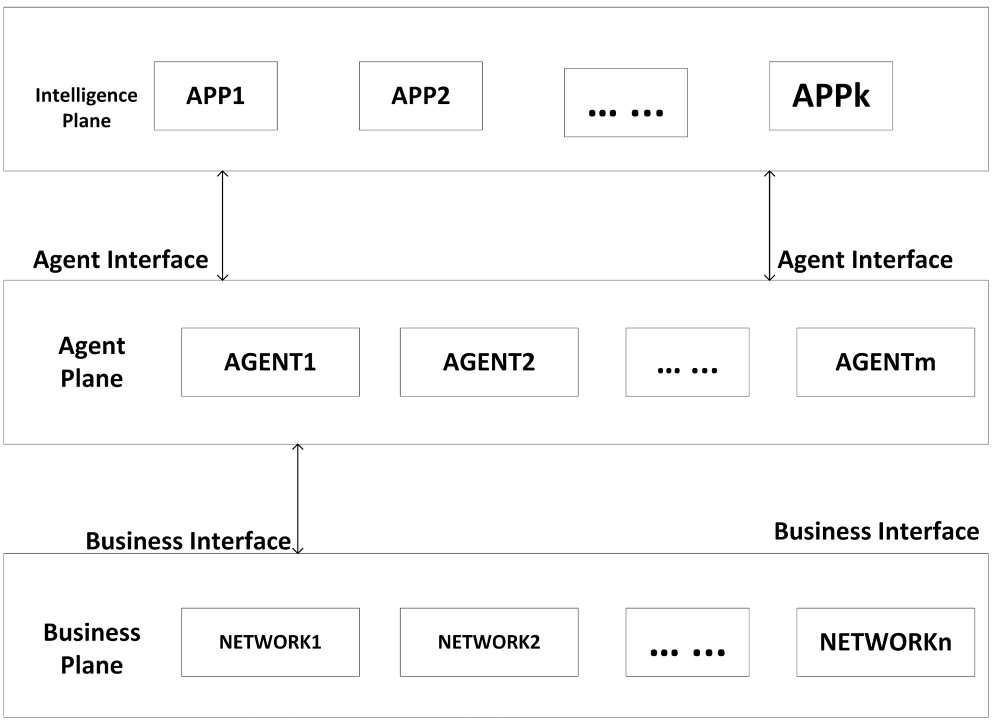

Figure 18. A system architecture of FINE [24].

The platform layer provides an intelligent plane for realising the intelligent logic of AI ability and behaviour. These include intelligent perception, machine mind and intelligent action. This layer helps in identifying and analysing predicted trends from the core layer. The application and terminal layer provide functionalities needed by the solution layer. Such functionalities includes load

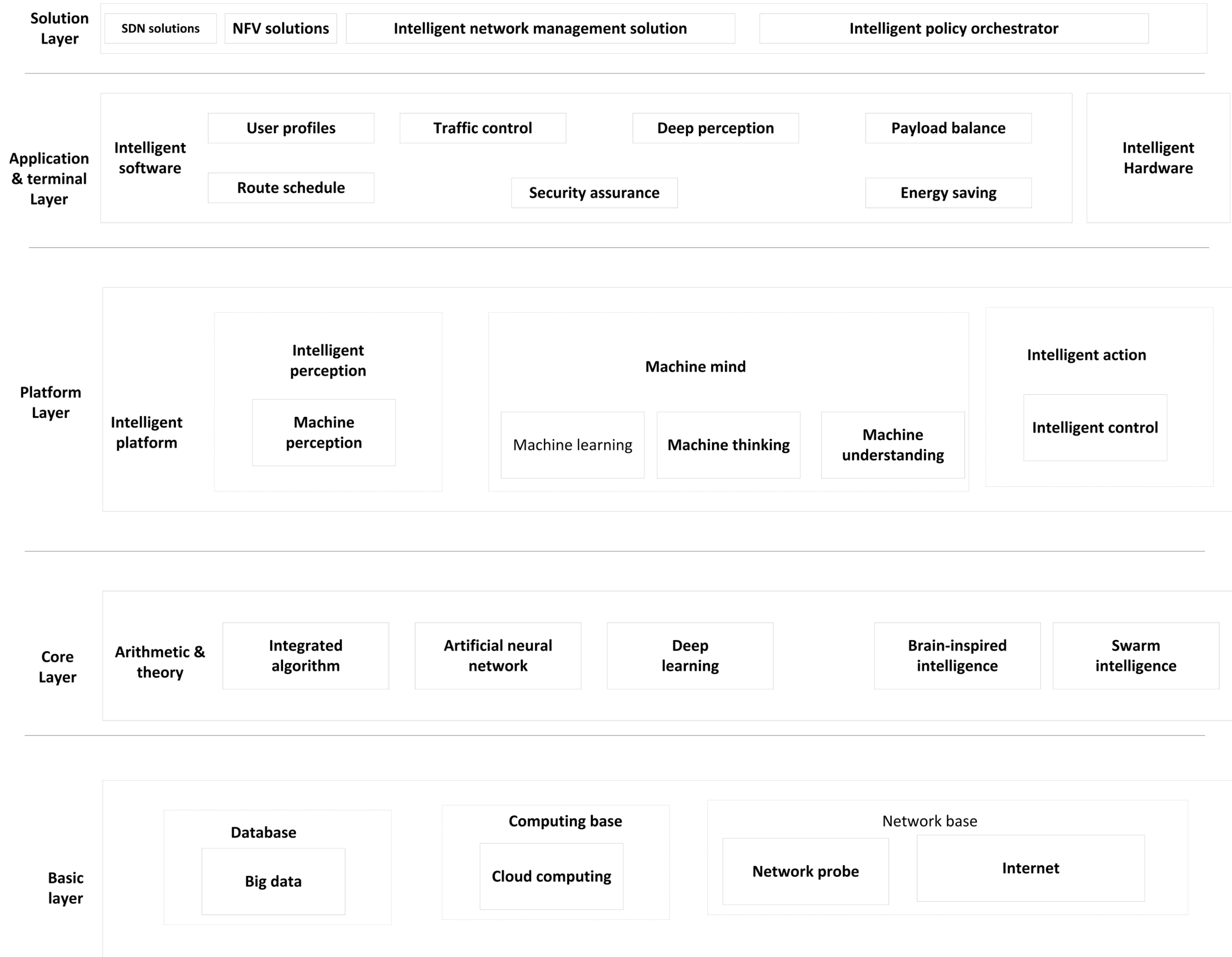

Figure 19. The Intelligence plane of the FINE framework [24].

balancing, security and energy saving. These realisations are identified in software and hardware. The solution layer is in charge of designing policies and related activities in managing the network.

The Agent Plane of the FINE framework serves as communicators amongst the different planes by sending intelligent control instructions. The Business Plane is in charge of executing services orchestrated by the intelligence plane. The Business Plane sends instructions to the Controller of SDN through Agents in the Agent Plane.

3.6. Big Data Analytics and Machine Learning Integrated in SDN and NFV

Le *et al.* [26] integrated machine learning algorithms at the forwarding layer of SDN to effectively cluster and forecast traffic behaviours of cells. The K-mean algorithm was used to cluster similar cells with the same traffic behaviour. A traffic forecasting model for each cluster using various ML algorithms was proposed to predict future traffic behaviours. As shown in **Figure 20**, the framework has an intelligent computing environment close to the RAN to support

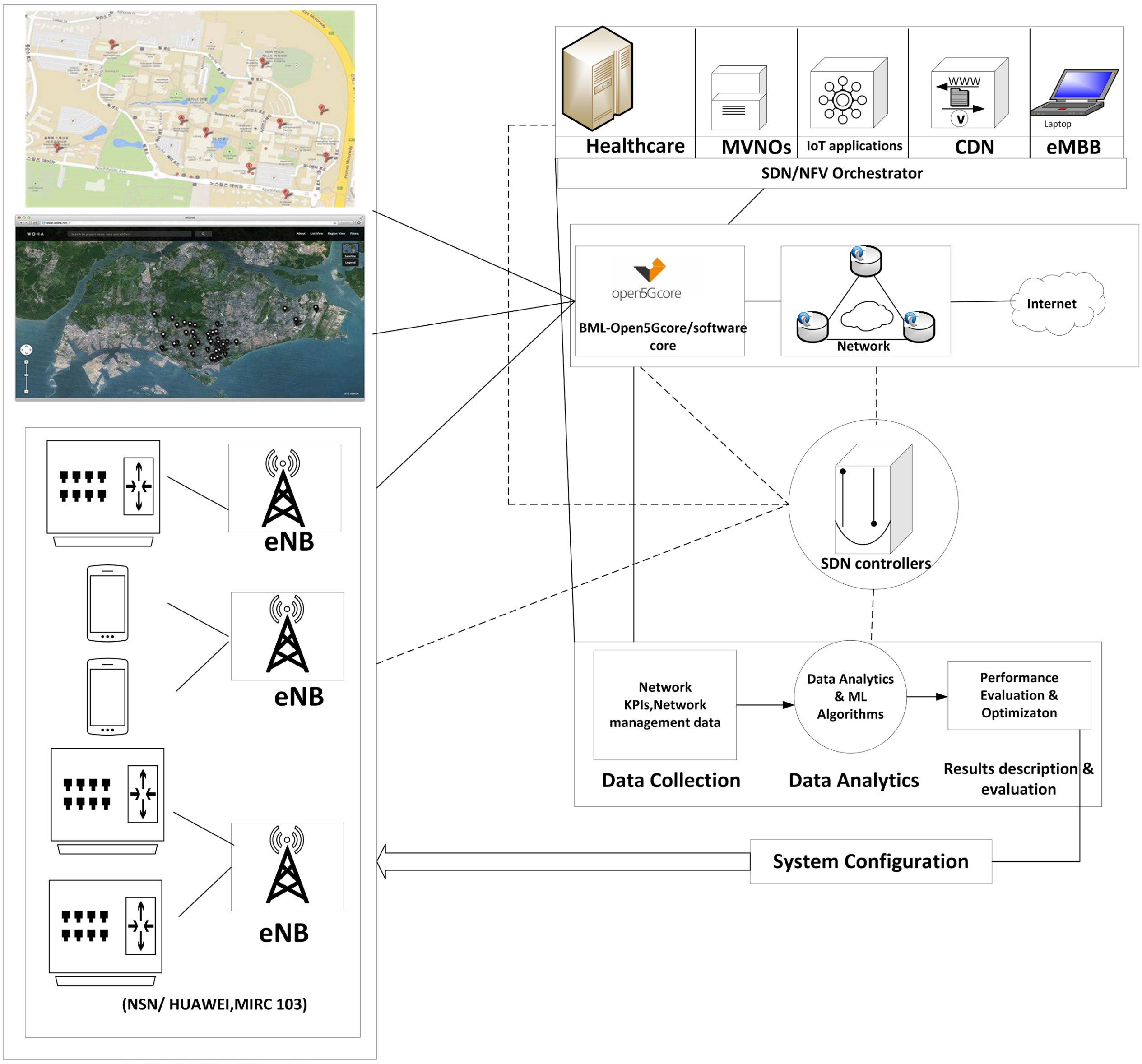

Figure 20. Big data analytics and machine learning ıntegrated in SDN and NFV [26].

Mobile Edge Computing (MEC). It is used to develop enhanced services such as Content Delivery Networks (CDN) and IoT applications.

The framework operates in building SON by:

- Collecting and storing mobile traffic and other information from all network sources.
- Using ML algorithms to extract and analyze the collected data to develop optimised models.
- Apply the models to configure the SDN controller and RANs components. Based on the model generated by the ML algorithms, the network's Key Performance Indicator (KPI) is matched with its requirements. If the network behavior achieves the expected performance, the new network parameters (NPs) will be used. Otherwise, relearn data based on new identified problems.

3.7. Artificial Intelligence Integrated Intrusion Detection Systems (IDS) for SDN

Li *et al.* [27] proposed an intelligent IDS with machine learning algorithms for SDN 5G architecture [28]. 5G-enabled IoT devices' surge in data traffic will lead to an uprising of novel attacks on networks. As shown in Figure 21, the architecture has three layers: Forwarding Layer, Management & Control Layer and Data & Intelligence Layer. The traffic monitoring and capturing are controlled by the Open Flow Entities (OFs), a component of the Forwarding Layer. The OFS, based on the instructions of the Controller, blocks malicious network flows

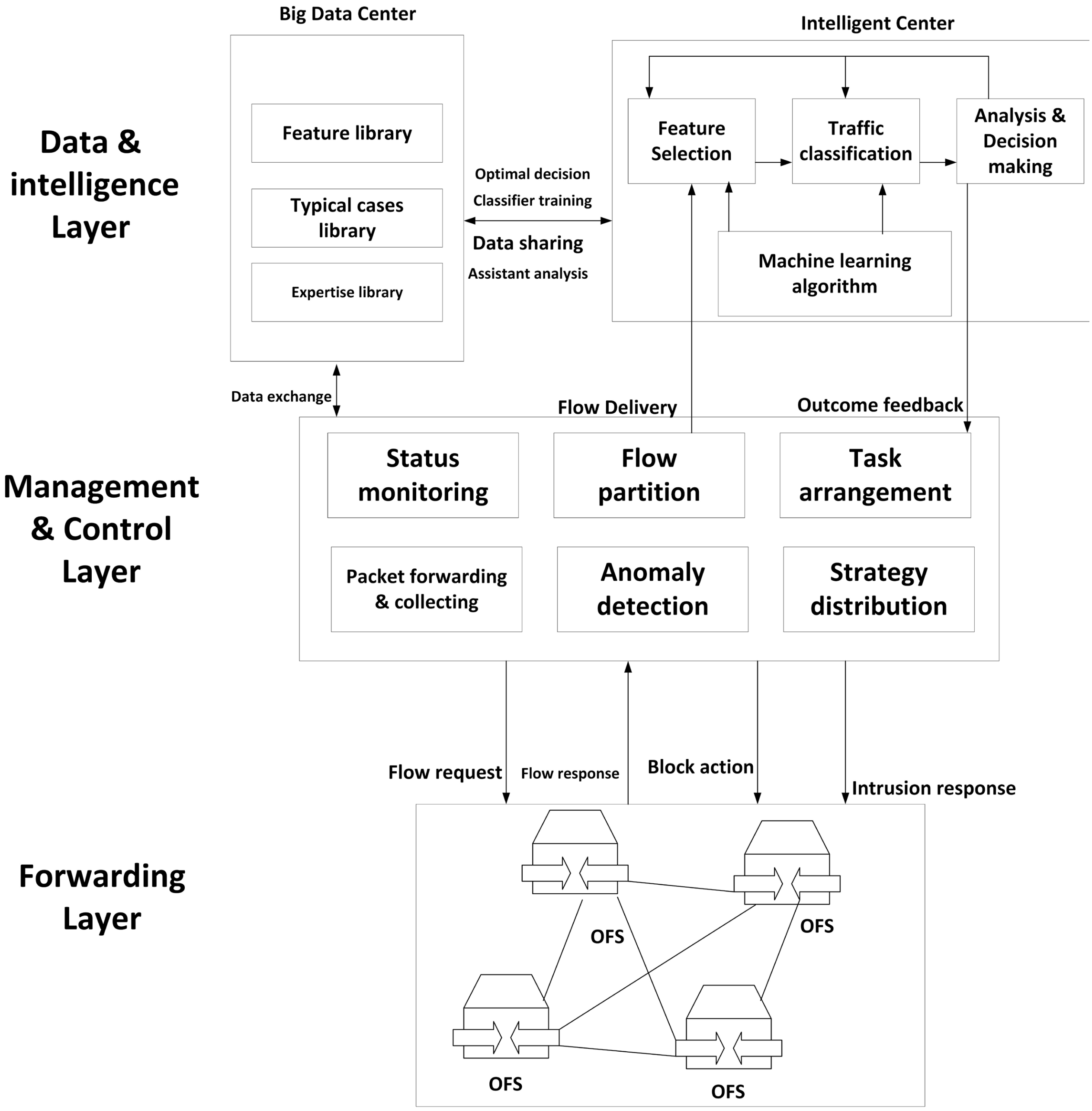

Figure 21. A machine learning IDS for software-defined 5G network [27].

uploaded to the Control Layer. A critical component of this architecture is the Data & Intelligence Layer. This intelligence centre makes further analysis and judgment of the network using machine learning algorithms.

The machine learning module of the Intelligent Center comprises Random Forest [29], K-means [30] and Adaboost [31] algorithms that enable the centre to perform feature selection and traffic classification. The feedback is forwarded to the Controller to make informed decisions. The informed decision is a comprehensive analysis determining whether the network is under attack. One other aspect of this architecture is the Big Data Center. This auxiliary module helps keep historical records and knowledge of intrusions to facilitate classifier training and decision-making. The data is periodically collected and updated.

A summary of each literature and technique is depicted in Table 1.

Table 1. Summary of AI-SON Techniques in 5G and SDN.

LITERATURE	5G	SDN	NFV	GENERIC AI TECHNIQUE	SPECIFIC AI TECHNIQUE
Pérez-Romero *et al.* [7]	✓			• Classification • Prediction	✖
Wang *et al.* [9]	✓			✖	✖
Monge *et al.* [10]	✓			• Pattern Recognition • Prediction	✖
Fu *et al.* [11]	✓			• Prediction • Clustering	✖
Wang *et al.* [12]	✓			• Prediction • Classification	✖
Zheng *et al.* [13]	✓			✖	✖
SELFNET [14]	✓			• Pattern Recognition • Prediction	✖
Sun *et al.* [19]	✓	✓		✖	✖
Sánchez *et al.* [21]	✓	✓		✖	✖
Bradai *et al.* [22]	✓	✓		✖	✖
Javaid *et al.* [23]	✓	✓		✖	✖
Xu *et al.* [24]	✓	✓	✓	• Artificial Neural Network • Swarm Intelligence • Brain-inspired Intelligence	✖
Le *et al.* [26]	✓	✓	✓	✖	✖
Li *et al.* [27]	✓	✓		• Classification • Clustering	• K-means • Adaboost • Random Forest

4. Open Issues

The research on intelligence model design in 5G and SDN has created gaps that need further analysis for an ideal model to be established. The intelligence design dynamics, especially in the SDN framework based on the extensive review, have exposed the following issues.

4.1. Feature Selection

Feature selection is an important concept in Machine Learning that hugely impacts the performance of a model [32] [33]. The best attributes that contribute most to a predicted variable's accuracy are maintained in feature selection. When irrelevant features are removed, overfitting [34] [35] [36] and model training time [37] [38] are significantly reduced. Most feature selection algorithms [39] [40] [41] handle large dimensionality and a huge number of instances that deals with data without class labels. MNOs have extremely heterogeneous data sources that require appropriate feature selection algorithms for accurate model design. The classifier built based on the selected features will increase accuracy in network load predictions and customer recommender systems.

4.2. Big Data Module

Big Data refers to a large, diverse collection of data that is huge in volume and grows exponentially with time [42] [43]. IoT devices are expected to generate a large amount of structured and unstructured data that must be analysed by the MNO [44]. Various devices create this heterogeneous data, leading to the four V's phenomenon, Volume, Velocity, Variety and Value of the data [45]. One aspect of the AI-SON model for 5G and SDN is the Big Data module, which helps in data acquisition, processing, storage and visualisation [46]. The layer in the SDN framework to deploy the Big Data module resulting in higher data efficiency and based on feature selection, is of concern. [24] deploys the Big Data module at the Application, Control and Infrastructure plane, whiles [22] proposed a knowledge layer on top of the Application layer for Big Data analytics. [29] proposed the Big Data module to be implemented at the Infrastructure plane. The Big Data module placement in the SDN framework for the highest level of data acquisition and linked to the Intelligence module is a research problem.

4.3. Machine Learning Algorithms

Artificial Intelligence algorithms are broad but generally categorised into Supervised Learning [47], Unsupervised Learning [48] and Reinforcement Learning [48]. Each machine learning technique achieves unique results based on the solution expected, the attributes and the dataset in building the machine learning model [49]. Classification [50] [51] [52] [53] and Regression [54] [55] [56] [57] [58] algorithms are suitable for Supervised Learning whiles Clustering [59] [60]

[61] [62] algorithms for Unsupervised Learning. Reinforcement learning algorithms [61] [62] are suitable for automation of systems. AI-SON model for 5G and SDN requires Machine Learning algorithms that adapt based on the dataset and the feature selection algorithm. The self-learning module implemented should select the best algorithms based on the accuracy of the predicted variables. There is a link between the attributes identified for transient load prediction from IoT-based devices and the selected algorithms in building a model by the MNOs.

4.4. Intelligence Module with Recommender Systems

The engine for AI-SON model for 5G and SDN is an intelligence module that predicts transient load, tracks patterns occurrence and provides recommender systems [63] to IoT-based clients. The intelligence module must provide QoS [64] to the MNO based on efficient predictor variables and QoE [64] to the IoT-based clients through Recommender Systems. In AI-SON model, building the intelligence module on either the client or server is a research issue [9] [13]. In AI-SON for SDN, the layer to build the intelligence module is another challenge [24] [26]. The location of the intelligence module and the big data module will affect the accuracy of predictions and the efficiency in resource utilisation by the MNOs and IoT-based clients. The design of the knowledge base that serves as feedback compared to new intelligence is a research challenge since the knowledge base needs some tested machine learning algorithms to determine the best knowledge catalogue [65] to compare with the new intelligence.

5. Conclusions

Several AI-SON integrated SDN has been reviewed in this study. The fast emergence of 5G networks coupled with a network architecture shift to SDN has necessitated a research direction in designing AI-SON models to cater to the exponential data growth. The AI-SON models identified from this review are suitable for 5G networks, and some are adapted to the SDN framework. Most of the researchers from this review created the Big Data and Intelligence modules in their frameworks and acknowledged the importance of A-I algorithms for efficient data analytics. The reviewed literature specific to the SDN framework identified the creation of the Big Data and Intelligence modules in either the Application plane or the Forwarding plane.

The review shows that most researchers have not proposed specific AI algorithms to test the Big Data gathered from internal or external sources. When done, this will help identify suitable machine learning algorithms for their models based on the data sources and feature selection criteria. In addition, cluster-intelligent recommender systems, when proposed, will help in Pattern Prediction (PP), Quality of Experience (QoE), Quality of Service (QoS) and Service-Level Agreement (SLA) analysis for end-users. This personalised user experience will gather predictive intelligence for IoT devices based on QoE and SLA.

AI-SON integrated SDN is vital in reducing the cost of 5G implementation to Mobile Network Operators (MNOs). The expected exponential growth in data will require the MNOs to either increase infrastructure or adopt advanced AI in network analysis for predictions.

Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

References

[1] Ericsson White Paper (2011) More than 50 Billion Connected Devices. https://vdna.be/publications/Wp-50-Billions.Pdf

[2] The 1000x Data Challenge. Qualcomm. https://www.qualcomm.com/content/dam/qcomm-martech/dm-assets/documents/Web_1000x_Tech_Exe_111913.pdf

[3] Celdrán, A.H., Pérez, M.G., Clemente, F.J.G. and Pérez, G.M. (2019) Towards the Autonomous Provision of Self-Protection Capabilities in 5G Networks. *Journal of Ambient Intelligence and Humanized Computing*, **10**, 4707-4720. https://doi.org/10.1007/s12652-018-0848-6

[4] Østerbø, O. and Grøndalen, O. (2012) Benefits of Self-Organising Networks (SON) for Mobile Operators. *Journal of Computer Networks and Communications*, 2012, Article ID: 862527. https://doi.org/10.1155/2012/862527

[5] METIS (2015) Mobile and Wireless Communications Enablers for the 2020 Information Society. https://cordis.europa.eu/project/id/671680

[6] Ericsson (2014) 5G Radio Access. *Ericsson Technology Review*, **6**, 1-8.

[7] Pérez-Romero, J., *et al.* (2015) Artificial Intelligence-Based 5G Network Capacity Planning and Operation. 2015 *International Symposium on Wireless Communication Systems*, Brussels, 25-28 August 2015, 246-250. https://doi.org/10.1109/ISWCS.2015.7454338

[8] Banerjee, A. (2014) Advanced Predictive Network Analytics: Optimise Your Network Investment and Transform Customer Experience. White Paper, Heavy Reading.

[9] Wang, Y., *et al.* (2017) A Data-Driven Architecture for Personalised QoE Management in 5G Wireless Networks. *IEEE Wireless Communications*, **24**, 102-110. https://doi.org/10.1109/MWC.2016.1500184WC

[10] Marco, M.S., Vidal, J.M. and Villalba, L.G. (2017) Reasoning and Knowledge Acquisition Framework for 5G Network Analytics. *Sensors*, **17**, Article 2405. https://doi.org/10.3390/s17102405

[11] Fu, Y., Wang, S., Wang, C.X., Hong, X. and McLaughlin, S. (2018) Artificial Intelligence to Manage Network Traffic of 5G Wireless Networks. *IEEE Network*, **32**, 58-64. https://doi.org/10.1109/MNET.2018.1800115

[12] Wang, J.B., *et al.* (2018) A Machine Learning Framework for Resource Allocation Assisted by Cloud Computing. *IEEE Network*, **32**, 144-151. https://doi.org/10.1109/MNET.2018.1700293

[13] Zheng, K., *et al.* (2016) Big Data-Driven Optimisation for Mobile Networks toward 5G. *IEEE Network*, **30**, 44-51. https://doi.org/10.1109/MNET.2016.7389830

[14] SELFNET Project (2020) Framework for Self-Organised Network Management in Virtualized and Software Defined Networks.

[15] Li, Y. and Chen, M. (2015) Software-Defined Network Function Virtualisation: A Survey. *IEEE Access*, **3**, 2542-2553. https://doi.org/10.1109/ACCESS.2015.2499271

[16] Casado, M., Koponen, T., Shenker, S. and Tootoonchian, A. (2012) Fabric: A Retrospective on Evolving SDN. *Association for Computing Machinery*, New York, NY, 85-90. https://doi.org/10.1145/2342441.2342459

[17] Foster, N., Harrison, R., Freedman, M.J., Monsanto, C., Rexford, J., Story, A. and Walker, D. (2011) Frenetic: A Network Programming Language. *ACM Sigplan Notices*, **46**, 279-291. https://doi.org/10.1145/2034574.2034812

[18] Bifulco, R., Canonico, R., Brunner, M., Hasselmeyer, P. and Mir, F. (2012) A Practical Experience in Designing an OpenFlow Controller. 2012 *European Workshop on Software Defined Networking*, Darmstadt, 25-26 October 2012, 61-66. https://doi.org/10.1109/EWSDN.2012.10

[19] Sun, S.L., *et al.* (2015) An Intelligent SDN Framework for 5G Heterogeneous Networks. *IEEE Communications Magazine*, **53**, 142-147. https://doi.org/10.1109/MCOM.2015.7321983

[20] Nunes, B.A.A., *et al.* (2014) A Survey of Software-Defined Networking: Past, Present, and Future of Programmable Networks. *IEEE Communications Surveys & Tutorials*, **16**, 1617-1634. https://doi.org/10.1109/SURV.2014.012214.00180

[21] José, S., *et al.* (2014) Softwarized 5G Networks Resiliency with Self-Healing. 1*st International Conference on* 5*G for Ubiquitous Connectivity*, Akaslompolo, 26-28 November 2014, 229-233.

[22] Abbas, B., *et al.* (2015) Cellular Software Defined Networking: A Framework. *IEEE Communications Magazine*, **53**, 36-43. https://doi.org/10.1109/MCOM.2015.7120043

[23] Nadeem, J., *et al.* (2018) Intelligence in IoT-Based 5G Networks: Opportunities and Challenges. *IEEE Communications Magazine*, **56**, 94-100. https://doi.org/10.1109/MCOM.2018.1800036

[24] Xu, G.B., Mu, Y.B. and Liu, J.L. (2017) Inclusion of Artificial Intelligence in Communication Networks and Services. *ICT Discoveries*, **1**, 1-6.

[25] Al-Hisnawi, M. and Ahmadi, M. (2016) Deep Packet Inspection Using Quotient Filter. *IEEE Communications Letters*, **20**, 2217-2220. https://doi.org/10.1109/LCOMM.2016.2601898

[26] Luong-Vy, L., *et al.* (2018) Applying Big Data, Machine Learning, and SDN/NFV to 5G Traffic Clustering, Forecasting, and Management. 2018 4*th IEEE Conference on Network Softwarization and Workshops*, Montreal, QC, 25-29 June 2018, 168-176.

[27] Li, J.Q., Zhao, Z.F. and Li, R.P. (2017) Machine Learning-Based IDS for Software-Defined 5G Network. *IET Networks*, **7**, 53-60. https://doi.org/10.1049/iet-net.2017.0212

[28] Li, R.P., *et al.* (2017) Intelligent 5G: When Cellular Networks Meet Artificial Intelligence. *IEEE Wireless Communications*, **24**, 175-183. https://doi.org/10.1109/MWC.2017.1600304WC

[29] Kamaldeep, S., *et al.* (2014) Big Data Analytics Framework for Peer-to-Peer Botnet Detection Using Random Forests. *Information Sciences*, **278**, 488-497. https://doi.org/10.1016/j.ins.2014.03.066

[30] David, A. and Vassilvitskii, S. (2006) K-Means++: The Advantages of Careful Seeding. *Proceedings of the Eighteenth Annual ACM-SIAM symposium on Discrete Algo-*

rithms, Stanford, 7-9 January 2007, 1027-1035.

[31] Trevor, H., *et al.* (2009) Multi-Class Adaboost. *Statistics and ITS Interface*, **2**, 349-360. https://doi.org/10.4310/SII.2009.v2.n3.a8

[32] Saeys, Y., Inza, I. and Larrañaga, P. (2007) A Review of Feature Selection Techniques in Bioinformatics. *Bioinformatics*, **23**, 2507-2517. https://doi.org/10.1093/bioinformatics/btm344

[33] Kira, K. and Rendell, L.A. (1992) A Practical Approach to Feature Selection. In: Sleeman, D. and Edwards, P., Eds., *Machine Learning Proceedings* 1992, Morgan Kaufmann, Burlington, 249-256. https://doi.org/10.1016/B978-1-55860-247-2.50037-1

[34] Pham, H.N.A. and Triantaphyllou, E. (2008) The Impact of Overfitting and Overgeneralisation on the Classification Accuracy in Data Mining. In: Maimon, O. and Rokach, L., Eds., *Soft Computing for Knowledge Discovery and Data Mining*, Springer, Boston, MA, 391-431. https://doi.org/10.1007/978-0-387-69935-6_16

[35] Ng, A.Y. (1997) Preventing Overfitting of Cross-Validation Data. *ICML*, **97**, 245-253.

[36] Reunanen, J. (2003) Overfitting in Making Comparisons between Variable Selection Methods. *Journal of Machine Learning Research*, **3**, 1371-1382.

[37] Ma, C. and Wang, J. (2018) Model Reduction with Memory and the Machine Learning of Dynamical Systems. *Communications in Computational Physics*, **25**, 947-962. https://doi.org/10.4208/cicp.OA-2018-0269

[38] Mai, L., Hong, C. and Costa, P. (2015) Optimising Network Performance in Distributed Machine Learning. 7*th USENIX Workshop on Hot Topics in Cloud Computing* (*HotCloud'*15), Santa Clara, 20-21 June 2016, 1-7.

[39] Langley, P. (1994) Selection of Relevant Features in Machine Learning. *Proceedings of the AAAI Fall Symposium on Relevance*, **184**, 245-271. https://doi.org/10.21236/ADA292575

[40] Dash, M. and Liu, H. (1997) Feature Selection for Classification. *Intelligent Data Analysis*, **1**, 131-156. https://doi.org/10.3233/IDA-1997-1302

[41] Hall, M.A. (1999) Correlation-Based Feature Selection for Machine Learning. Ph.D. Thesis, The University of Waikato, Waikato.

[42] Kambatla, K., Kollias, G., Kumar, V. and Grama, A. (2014) Trends in Big Data Analytics. *Journal of Parallel and Distributed Computing*, **74**, 2561-2573. https://doi.org/10.1016/j.jpdc.2014.01.003

[43] LaValle, S., Lesser, E., Shockley, R., Hopkins, M.S. and Kruschwitz, N. (2011) Big Data, Analytics and the Path from Insights to Value. *MIT Sloan Management Review*, **52**, 21-32.

[44] Mourtzis, D., Vlachou, E. and Milas, N. (2016) Industrial Big Data as a Result of IoT Adoption in Manufacturing. *Procedia CIRP*, **55**, 290-295. https://doi.org/10.1016/j.procir.2016.07.038

[45] Malek, Y.N., Kharbouch, A., El Khoukhi, H., Bakhouya, M., De Florio, V., El Ouadghiri, D. and Blondia, C. (2017) On the Use of IoT and Big Data Technologies for Real-Time Monitoring and Data Processing. *Procedia Computer Science*, **113**, 429-434. https://doi.org/10.1016/j.procs.2017.08.281

[46] Mariscal, G., Marban, O. and Fernandez, C. (2010) A Survey of Data Mining and Knowledge Discovery Process Models and Methodologies. *The Knowledge Engineering Review*, **25**, 137-166. https://doi.org/10.1017/S0269888910000032

[47] Caruana, R. and Niculescu-Mizil, A. (2006) An Empirical Comparison of Super-

vised Learning Algorithms. *Proceedings of the 23rd International Conference on Machine Learning*, Pittsburgh, 25-29 June 2006, 161-168. https://doi.org/10.1145/1143844.1143865

[48] Ghahramani, Z. (2003) Unsupervised Learning. In: Bousquet, O., von Luxburg, U. and Rätsch, G., Eds., *Summer School on Machine Learning*, Springer, Berlin, 72-112. https://doi.org/10.1007/978-3-540-28650-9_5

[49] Dey, A. (2016) Machine Learning Algorithms: A Review. *International Journal of Computer Science and Information Technologies*, **7**, 1174-1179.

[50] Suykens, J.A. and Vandewalle, J. (1999) Least Squares Support Vector Machine Classifiers. *Neural Processing Letters*, **9**, 293-300. https://doi.org/10.1023/A:1018628609742

[51] Friedman, J.H. (1989) Regularised Discriminant Analysis. *Journal of the American Statistical Association*, **84**, 165-175. https://doi.org/10.1080/01621459.1989.10478752

[52] Zhang, H. (2004) The Optimality of Naive Bayes. The Florida AI Research Society, Florida.

[53] Keller, J.M., Gray, M.R. and Givens, J.A. (1985) A Fuzzy K-Nearest Neighbor Algorithm. *IEEE Transactions on Systems, Man, and Cybernetics*, **SMC-15**, 580-585. https://doi.org/10.1109/TSMC.1985.6313426

[54] Guisan, A., Weiss, S.B. and Weiss, A.D. (1999) GLM versus CCA Spatial Modeling of Plant Species Distribution. *Plant Ecology*, **143**, 107-122. https://doi.org/10.1023/A:1009841519580

[55] Nguyen-Tuong, D., Peters, J.R. and Seeger, M. (2009) Local Gaussian Process Regression for Real Time Online Model Learning. In: *Advances in Neural Information Processing Systems*, MIT Press, Cambridge, 1193-1200.

[56] Dietterich, T.G. (2000) Ensemble Methods in Machine Learning. In: *Multiple Classifier Systems. MCS* 2000. *Lecture Notes in Computer Science*, Vol. 1857, Springer, Berlin, 1-15. https://doi.org/10.1007/3-540-45014-9_1

[57] Safavian, S.R. and Landgrebe, D. (1991) A Survey of Decision Tree Classifier Methodology. *IEEE Transactions on Systems, Man, and Cybernetics*, **21**, 660-674. https://doi.org/10.1109/21.97458

[58] Sudheer, K.P., Gosain, A.K., Mohana, R.D. and Saheb, S.M. (2002) Modelling Evaporation Using an Artificial Neural Network Algorithm. *Hydrological Processes*, **16**, 3189-3202. https://doi.org/10.1002/hyp.1096

[59] Alsabti, K., Ranka, S. and Singh, V. (1997) An Efficient K-Means Clustering Algorithm: Analysis and Implementation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **24**, 881-892.

[60] Johnson, S.C. (1967) Hierarchical Clustering Schemes. *Psychometrika*, **32**, 241-254. https://doi.org/10.1007/BF02289588

[61] Reynolds, D.A. (2009) Gaussian Mixture Models. In: Li, S.Z. and Jain, A., Eds., *Encyclopedia of Biometrics*, Springer, Boston, MA, 659-663. https://doi.org/10.1007/978-0-387-73003-5_196

[62] Beal, M.J., Ghahramani, Z. and Rasmussen, C.E. (2002) The Infinite Hidden Markov Model. In: Dietterich, T.G., Becker, S. and Ghahramani, Z., Eds., *Advances in Neural Information Processing Systems*, MIT Press, Cambridge, 577-584.

[63] Rashid, A.M., Karypis, G. and Riedl, J. (2008) Learning Preferences of New Users in Recommender Systems: An Information Theoretic Approach. *ACM SIGKDD Explorations Newsletter*, **10**, 90-100. https://doi.org/10.1145/1540276.1540302

[64] Karakus, M. and Durresi, A. (2017) Quality of Service (QoS) in Software Defined Networking (SDN): A Survey. *Journal of Network and Computer Applications*, **80**, 200-218. https://doi.org/10.1016/j.jnca.2016.12.019

[65] Huang, X. and Jensen, J.R. (1997) A Machine-Learning Approach to Automated Knowledge-Base Building for Remote Sensing Image Analysis with GIS Data. *Photogrammetric Engineering and Remote Sensing*, **63**, 1185-1193.