ITUPublications

International Telecommunication Union
Radiocommunication Sector

Report ITU-R SM.2542-0

(06/2024)

SM Series: Spectrum management

Next generation spectrum monitoring – proactive, autonomous and data-driven

Foreword

The role of the Radiocommunication Sector is to ensure the rational, equitable, efficient and economical use of the radio-frequency spectrum by all radiocommunication services, including satellite services, and carry out studies without limit of frequency range on the basis of which Recommendations are adopted.

The regulatory and policy functions of the Radiocommunication Sector are performed by World and Regional Radiocommunication Conferences and Radiocommunication Assemblies supported by Study Groups.

Policy on Intellectual Property Right (IPR)

ITU-R policy on IPR is described in the Common Patent Policy for ITU-T/ITU-R/ISO/IEC referenced in Resolution ITU-R 1. Forms to be used for the submission of patent statements and licensing declarations by patent holders are available from http://www.itu.int/ITU-R/go/patents/en where the Guidelines for Implementation of the Common Patent Policy for ITU-T/ITU-R/ISO/IEC and the ITU-R patent information database can also be found.

Series of ITU-R Reports

(Also available online at https://www.itu.int/publ/R-REP/en)

Series	Title
BO	Satellite delivery
BR	Recording for production, archival and play-out; film for television
BS	Broadcasting service (sound)
BT	Broadcasting service (television)
F	Fixed service
M	Mobile, radiodetermination, amateur and related satellite services
P	Radiowave propagation
RA	Radio astronomy
RS	Remote sensing systems
S	Fixed-satellite service
SA	Space applications and meteorology
SF	Frequency sharing and coordination between fixed-satellite and fixed service systems
SM	**Spectrum management**
TF	Time signals and frequency standards emissions

***Note**: This ITU-R Report was approved in English by the Study Group under the procedure detailed in Resolution ITU-R 1.*

Electronic Publication
Geneva, 2024

REPORT ITU-R SM.2542-0

Next generation spectrum monitoring – proactive, autonomous and data-driven

(2024)

TABLE OF CONTENTS

Page

Scope

This Report applies new trends in data sciences including artificial intelligence and big data technologies to the automation of spectrum monitoring. The Report summarizes the current reactive and ad hoc spectrum management approaches, and then details next-generation approaches that are proactive, autonomous and data-driven, using artificial intelligence and big data. The Report gives several implementation examples that illustrate the power of these techniques.

1 Terms, definitions and abbreviations

1.1 Big data

Recommendation ITU-T Y.3600 – Big data – Cloud computing based requirements and capabilities, defines big data as "a paradigm for enabling the collection, storage, management, analysis and visualization, potentially under real-time constraints, of extensive datasets with heterogeneous characteristics." Examples of datasets characteristics include high-volume, high-velocity, high-variety, etc. This Report discusses generation of spectrum big data from densely distributed monitoring networks that will require both local and central processing to route, store and display the resulting information. Because of the network and cloud-computing requirements for such a deployment, ITU publications dealing with information infrastructure are cited here.

1.2 Artificial intelligence

Recommendation ITU-T M.3080 – Framework of artificial intelligence enhanced telecom operation and management (AITOM), defines artificial intelligence (AI) as a computerized system that uses cognition to understand information and solve problems. ISO/IEC 2382-28 defines AI as "an interdisciplinary field, usually regarded as a branch of computer science, dealing with models and systems for the performance of functions generally associated with human intelligence, such as reasoning and learning". In computer science, AI research is defined as the study of "intelligent agents": any device that perceives its environment and takes actions to achieve its goals. This includes pattern recognition, the application of machine learning and related techniques. Artificial-intelligence is the whole idea and concept of machines being able to carry out tasks in a way that mimics human intelligence and would be considered "smart". Further studies in this respect are invited.

1.3 Machine learning

Recommendation ITU-T Y.3172 – Architectural framework for machine learning in future networks including IMT-2020, defines machine learning (ML) as processes that enable computational systems to understand data and gain knowledge from it without necessarily being explicitly programmed. Supervised machine learning and unsupervised machine learning are two examples of machine learning types.

1.4 Radio frequency machine learning

"… recent research has shown deep machine learning to be an enabling technology for cognitive radio applications as well as a useful tool for supplementing expertly defined algorithms for spectrum sensing applications such as signal detection, estimation and classification (termed here as Radio Frequency Machine Learning, or RFML). A major driver for the usage of deep machine learning in the context of wireless communications is that little, to no, a priori knowledge of the intended spectral environment is required, given that there is an abundance of representative data to facilitate training and evaluation." [1]

1.5 Abbreviations

AI	Artificial intelligence
AOA	Angle of arrival
API	Application programming interface
BI	Business intelligence
DB	Database
DF	Direction finder(ing)
DMR	Digital mobile radio
FFT	Fast Fourier transform
GUI	Graphical user interface
IoT	Internet of things
I/Q	In-phase/quadrature phase
KPI	Key performance indicator
LAN	Local area network
LTE	Long term evolution
ML	Machine learning
PU	Primary user
QoE	Quality of experience
QoS	Quality of service
RBW	Resolution bandwidth
RF	Radio frequency
RFML	Radio frequency machine learning
RSS	Received signal strength
SQL	Structured query language
TDOA	Time-difference-of-arrival
UAV	Unmanned aerial vehicle
WCDMA	Wideband code division multiple access

2 Introduction

The radio spectrum is a limited resource and the increase in demand for spectrum for emerging new radio technologies remains unabated. Continuous advancements in wireless communication technologies have resulted in new standards with:

– Higher carrier frequencies with diverse band assignments;

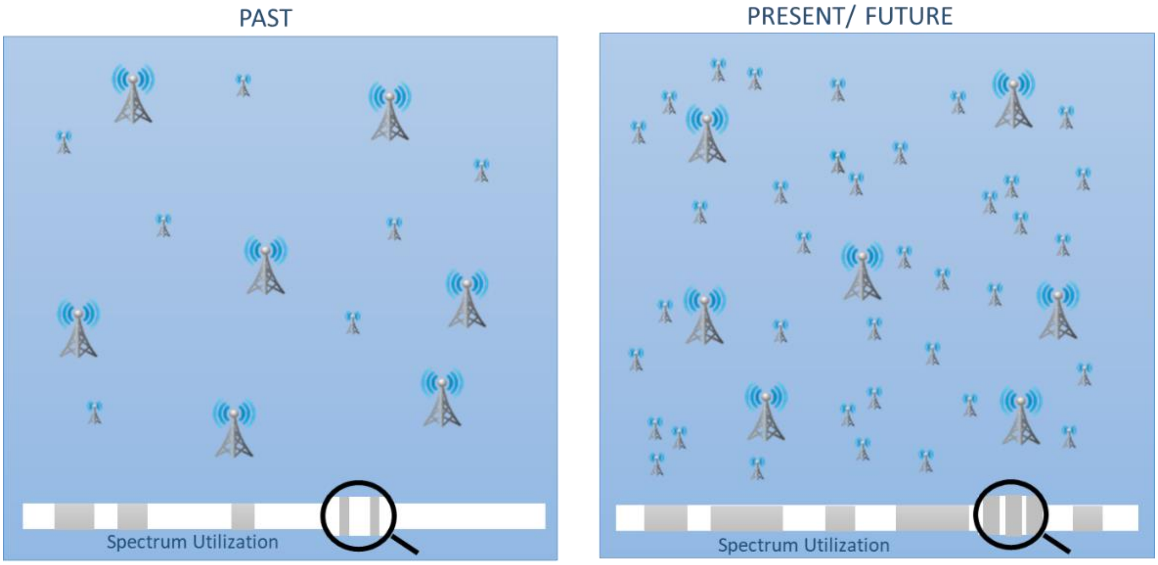
– Wider modulation bandwidth;

– Lower transmitter power levels to increase system capacity through frequency reuse.

These technology trends result in the deployment of radio communication systems that have a high number of lower power transmitters to provide coverage in a given geographic area. Modern spectrum monitoring strategies need to account for these types of emitter deployments, while still being able to monitor ‘traditional’ emitters such as FM and TV broadcast or air traffic communications based on traditional analogue AM transmissions.

FIGURE 1

Trend towards high-density of low-power emitters



Given these conditions, spectrum monitoring has become a complex task, requiring fine grained and dense spectrum usage data across frequency, time and location. Traditional approaches to spectrum monitoring are not always able to provide the density of data required to manage today’s spectrum.

Many national authorities maintain specialized services for the systematic monitoring of radio emissions to create knowledge about spectrum use and to provide feedback to the regulatory process and if needed, enforcement procedures. These services create occupancy data for the assessment of utilization [2] to efficiently manage the spectrum that forms the basis for many spectrum management tasks, from maximizing license revenues to verifying compliance and resolving interference problems. At the same time, the rapid expansion in wireless communications, both in the number of transmitters and in standards, make measuring and tracking changes in spectrum utilization more challenging.

Traditionally, spectrum monitoring relies on the skills of trained operators, using tools such as spectrograms, power spectral density and demodulation, to identify emitters, locate violators, and report on spectrum usage characteristics. On the other hand, while spectrum engineering tools such as coverage maps showing propagation models and effects offer visibility into spectrum usage, they rely on an individual’s expertise and/or tacit knowledge to analyse spectrum monitoring data and take the appropriate actions. These skills are hard to replicate or scale and limit the ability of administrations to face ever growing challenges to spectrum management.

On the other hand, future data-driven spectrum monitoring aims at less human intervention and more autonomous surveillance. For instance, data-driven spectrum monitoring can detect specific signals even in the case of a high-density of emitters. It can also collect large volumes of spectrum data autonomously. The data-driven spectrum monitoring can link between previous and future data due to accumulated database giving the potential to increase the efficiency of spectrum supervisors. Advanced spectrum monitoring systems can help administrations to manage abnormal spectrum events more easily and, depending on the network density, with a high coverage area.

Data-driven spectrum monitoring can be used to identify trends and anticipate future events by analysing the pattern of signals. This enables the monitoring service to prepare accordingly.

Spectrum allocation practices traditionally have embraced "Static Allocations" where a license is issued to a primary user exclusively on a long-term basis over large geographical regions. With the exception of some services e.g. the broadcast services, it is rare that licensees make full use of the frequencies assigned to them. In many cases, some portion of the allocated band(s) may remain underutilized. Data-driven spectrum monitoring may enable more efficient spectrum management practices.

To improve spectrum management practices, spectrum utilization needs to be accurately characterized at the national, regional, urban and local level. This drives the need to identify available spectrum across time and location in the area of interest. Available spectrum may also be referred to as spectrum holes – spatial and temporal. The temporal spectrum hole appears in a time interval when there is no Primary User transmission. The spatial spectrum hole appears in a location or area where no Primary User transmissions are detected.

3 Distributed spectrum monitoring

Next Generation Spectrum Management takes advantage of and aggregates the capacity of underutilized frequency bands employing an integrated approach that leverages new advances in radio frequency (RF) sensing, distributed and mobile sensing that enable geolocation measurements, centralized databases, analytics and ML to increase capacity, improve coverage and mitigate interference.

An autonomous approach to spectrum monitoring can adapt the monitoring tasks to changing conditions and directly answer meaningful queries through dashboards that may provide insight into radio spectrum KPIs, such as utilization, efficiency, allocation, users, and available spectrum by location, time and frequency band. An example of next generation distributed spectrum monitoring is given in Annex 1, describing the experience of Korea (Republic of) with the use of these techniques for spectrum monitoring.

3.1 Elements of a distributed spectrum monitoring system

A coordinated and concentrated effort to collect large-scale spectrum allocation and usage information at a micro level requires many geographically distributed spectrum monitoring sensors that are adaptable to changing operational environments and technical requirements.

Managing a network of distributed sensors, and the voluminous data they generate requires fusing data from multiple sensors and selecting the optimal dataset to drive decisions. This necessitates intelligent algorithms that can analyse the data and apply domain knowledge to generate relevant information for various stakeholders.

A fundamental task in spectrum monitoring is to identify specific usage patterns in the various dimensions (time, frequency, location, direction, etc.) to reveal available spectrum. Building detailed spectrum occupancy maps at the national, regional, urban and local level requires continuous spectrum monitoring, and is challenging and time consuming.

Based on the above, the following elements can form a scalable, dynamic spectrum monitoring system providing an economical national solution operating down to the local level, and over a large range of frequencies:

a) Fixed traditional monitoring systems will continue to be inevitable for monitoring, especially when large antennas are required. Examples for that apply to the monitoring at HF, VHF and UHF frequency bands or high sensitivity satellite monitoring stations. Such stations may also be useful to make DF measurements for interference resolution and identification of illegal transmitters. It has to be noted that this concept does not address organizational situations in case of network failure, i.e. the operation of a network independent manned monitoring station;

b) Lower cost small form factor RF sensor systems may be used for long term measurement, such as seasonal occupancy evaluation or to evaluate specific interference cases of intermittent weak signals that may require long term measurements;

c) Mobile and portable monitoring systems can be used to achieve a high spatial resolution for dynamic spectrum management which fixed stations cannot provide. Annex 2 to this Report provides examples of mobile and portable monitoring solutions to address localized spectrum measurements of field strength, coverage and detection of interference.

The large amount of monitoring data reported back to the control centre will add to big data management issues, making centralized databases and analytical tools with real-time dashboards critical.

Displays giving generalized locations of multiple emitters based on the license database, as well as DF and RSS measurements are useful in evaluating emitter density, spectrum occupancy and spectrum holes at the local level. Determining the precise location of interfering emitters will require TDOA, AOA and Hybrid geolocation techniques. Due to the low power of many emitters, mobile and distributed small sensor monitoring/geolocation systems will become more important. Additionally, these systems will need to detect and locate emitters in the higher frequency ranges.

3.1.1 Local/edge processing at monitoring receivers

The goal of local processing is to transform the raw wideband data created by the monitoring receiver into an aggregated and transportable form of information and moving spectral intelligence closer to the edge of the network. Having processing and storage at the local level results in faster analysis, comparative processing of spectrum snapshots, and real-time machine decision making. The use of signal processing, enabled by Digital Down Converters (DDCs), increases the visibility and accuracy of the signalling environment and changes within that environment. The local processing further reduces the burden on the communication network that ties together the receivers, central processors and data storage without losing important data and analysis results.

Artificial Intelligence (AI), particularly in the form of Radio Frequency Machine Learning (RFML), may become a necessity to process the large volume of real-time monitoring data across many locations. Sensors with local processors and RFML capabilities will be needed to reduce the amount of network backhaul data (refer to § 5). Less-capable secondary sensors can be used to extend geographic coverage in the most economical way.

Some I/Q time-series data will need to be recorded locally and transferred to the control centre for processing of unknown signals where it could be used to train a neural network to recognize the signal across the entire monitoring network.

3.1.2 Data security

Given the amount of data that may be collected and transferred to a control centre, it may be advisable to implement appropriate data security measures, which may involve data encryption, or other data security measures. This topic is addressed in detail in Recommendation ITU-T X.1601 – Security framework for cloud computing. New technologies such as Blockchain may be relevant to such networks. Blockchain (e.g. alliance chain mode) is a distributed ledger technology developed through the integration and innovation of computer technologies such as distributed data storage, peer-to-peer transmission, consensus mechanisms, encryption algorithms and smart contracts. It has the characteristics of transparency, traceability and tamper resistance.

3.2 Big data challenges of a distributed spectrum monitoring system

Deploying high numbers of distributed RF sensors and receivers to perform continuous monitoring across wide segments of the spectrum creates network capacity challenges which may result in high costs associated with maintaining network connections to the processing centre(s). The volume of monitoring data is directly related to the type of data being sent by the RF sensors – time-series data, FFT-based spectral data, or RF parametric data (described below). Each type of data is needed for different measurements typical to monitoring tasks. For example, signal classification and technical identification often requires I/Q time-series data whereas occupancy measurements can be done with either spectral data or RF parametric data. I/Q time-series data is the heaviest form of monitoring data needing the highest level of network capacity to transfer a specific bandwidth as compared with spectrum or RF parametric data.

3.2.1 Monitoring data types and impact on network bandwidth

As an example, performing continuous streaming of I/Q time-series data at 20 MHz acquisition bandwidth requires 1.25 (filter shape) * 20 MHz (bandwidth) * 2 (I and Q) * 16 bits (I or Q) = 800 Mbit/s data rate to be sustained, close to the speed of Gigabit LAN. A further constraint is that I/Q data capture must be continuous, so that the original modulation is not affected which is essential when classifying the signal or performing technical identification. One hour of continuous I/Q time-series streaming produces about 360 GB of monitoring data.

By comparison, observation of a continuous spectral streaming of a 20 MHz bandwidth using trace processing to reduce time resolution and assuming a 16k-point FFT roughly produces a 10 Mbit/s data rate. One hour of continuous FFT-based spectrum streaming produces considerably less data than I/Q streaming, roughly 4.5 GB. Reducing the spectrum update rate to 1 scan/s reduces the data rate to about 24 kB/s. In this case, one hour of 1 scan/s monitoring produces 86 MB. Transfer and storage of raw spectral data requires significant bandwidth and storage media.

RF parametric data is the result of processing performed on the spectral data from an RF sensor or receiver. A level or other form of threshold is applied to the spectral data and each energy exceeding the threshold has basic external parameters extracted and stored and/or accumulated for a period of time. This form of energy detection can benefit occupancy measurements and does not transfer or process noise (or signal) energy below the energy threshold (see Fig. 2). Parameters that can be extracted include:

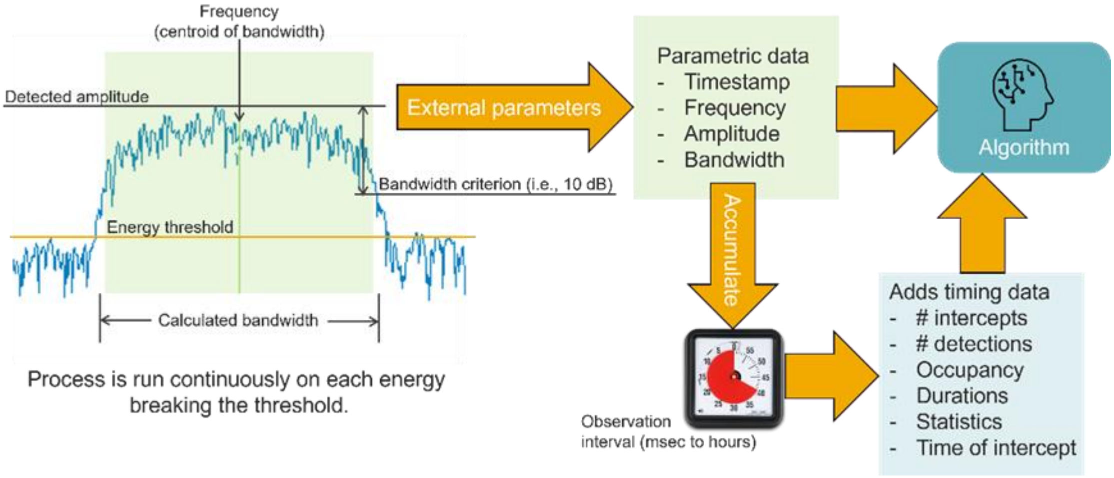
- centre frequency;
- bandwidth;
- amplitude/field strength;
- time stamp.

If an observation interval is applied to the data collection, additional parameters can be computed and stored. These are each computed within the specified interval:

- Percent occupancy;
- time of first and last intercept;
- duration (minimum, maximum, average);
- number of detections;
- bandwidth statistics (minimum, maximum, average);
- amplitude/field strength statistics (minimum, maximum, average).

FIGURE 2

Development of RF parametric data



Using an observation interval allows the spectral sweep to update in real time, while the data reporting is lessened to one record per detected energy per each observation interval.

This process improves the data efficiency of spectrum monitoring converting raw spectral data into textual data that can be transferred or stored for further processing, analysis, and provide well formatted data to ML algorithms.

To create FFT-spectral or RF parametric data, a monitoring system must have the capability to digitize the receiver output (FFT samples) in real-time and incorporate local processing. Going further, local intelligence to identify anomalies and access to the license database increases the sophistication and cost of the monitoring station but lowers the cost of the network traffic. Ideally, an RF monitoring station has the ability to produce all necessary forms of data to support different monitoring tasks. As monitoring systems become more densely deployed to detect low-level wireless services and localized interference, more processing aided by ML algorithms will be needed to reduce the network data as much as possible.

3.2.2 ML algorithms applied to spectrum monitoring

There is a large volume of technical papers and presentations on ML based spectrum monitoring. This Report provides a basic summary of the developments and applications. The first step in the process of development may be to discuss the purpose of effort to develop ML and what benefits will be obtained. Typical spectrum monitoring tasks that support decisions related to management include making frequency allocations, detecting and classifying interference, determining license compliance, and predictive maintenance (possibly for the case of terrestrial or satellite downlink monitoring).

As the concept of dynamic spectrum sharing develops and is realized in some cognitive networks (e.g. IoT), making frequency allocations in real-time, autonomously, based on deep learning of the local spectrum must be enabled by extensive monitoring of the bands – and more specifically – occupancy data [3]. Granular occupancy data of the frequencies to be shared in the appropriate locations will be essential to training the ML algorithms to realize intelligent allocations. Implementing this capability may require a combination of classification and regression test algorithms.

For the detection and classification of interference, an ML algorithm will need to observe and learn the local spectrum. Training data will need to represent normal signal energy and then use a form of anomaly detection to detect interference. Signal classification will also likely be important for this task and need to make use of neural networks – requiring I/Q time-series data from the monitoring receivers. See § 5.

License compliance and predictive maintenance applications will rely heavily on either spectral or RF parametric data (centre frequency, bandwidth, field strength) observed and learned over long periods of time and processed by a regression algorithm.

3.2.3 Common types of ML algorithms

There are four basic forms of ML that may be applied to spectrum monitoring: classification, regression test, anomaly detection, and clustering. These are described briefly.

– Classification answers the question, “Does this data have the attributes that belong with this group or that?” The output is a category rather than a value. The decision is based on comparison to the training data. This may be applied to occupancy data in support of autonomous frequency selection and allocation.

– Regression attempts to predict the relationship between independent and dependent datasets, looking for a line fit of the data representing a form of correlation. The output is typically a value between −1 and +1 where 0 represents no correlation, +1 positive correlation, and −1 negative correlation. Regression testing operating on FFT-spectrum or RF parametric data may be useful for predictive maintenance applied to a satellite earth station by closely monitoring RSS on a downlink and taking into consideration local environmental data on humidity, precipitation, temperature, and wind. It is important to consider that the RF parameters of an emission measured by a monitoring station may fluctuate due to changes in the local environment, traffic conditions, propagation changes caused by humidity or temporary blockage of the signal.

– Anomaly detection looks for outliers in the dataset. Traditional spectrum monitoring software has relied on rule-based anomaly detection using setpoints on key parameters to trigger an alert or take some action. An ML-based anomaly detector develops patterns in the training data and attempts to fit new observations into one of those patterns. Data that does not fit is flagged as a potential anomaly. Anomaly detection operating on spectral or RF parametric data may be useful for detection of interference or the emergence of an unknown emission.

– Clustering causes the input, which is usually raw unlabelled data, to be sorted into collections based on the similarity of parametric qualities describing the data. This may be useful when applied to detection and classification of novel signals where the frequency plan is unknown.

These algorithms and applications are only examples of what is available today commercially in spectrum monitoring solutions. Some current software solutions provide high accuracy signal classification of modern services such as 4G LTE, Wi-Fi, and 5G using I/Q time-series data processed by a neural network.

As monitoring stations may require more advanced processing, trade-offs between cost, sophistication, density of sensors, and network backhaul capacity will need to be evaluated prior to specifying a solution.

3.2.4 Metadata and time synchronization

Distributed monitoring systems likely need to embrace a store-forward architecture to account for network latencies and disruptions. Store-forward is a data communication technique in which a message transmitted from a source node is stored at an intermediary device before being forwarded to the destination node.

In a typical deployment, each monitoring site detects signals (frequency bands) of interest, performs some processing and generates detailed characteristics on the spectrum. As these monitoring nodes work independently, comparing data from different sensors without a reference base creates challenges in gaining valuable insights. Therefore, it is essential to have metadata on site readings accompanied with a precision time stamp so that monitoring sensors can be synchronized to undertake coordinated spectrum measurement surveys.

3.2.5 Sensor data format

There are many approaches to spectrum measurement resulting in many kinds of data, the data needs to be "graded" or "classified" by its characteristics. Examples of this might include:

a) RF background noise;

b) Energy or power spectral density;

c) I/Q samples or FFT samples;

d) Decimated I/Q samples;

e) Spectrum occupancy data;

f) Geolocation data;

g) Level of metadata supporting the measurement data.

3.2.6 Heterogeneity in spectrum monitoring equipment

A distributed spectrum monitoring system will use spectrum monitoring sensors with distinctive characteristics, operating parameters often from multiple vendors. A key challenge is to support different measurement equipment that support different communication APIs, sampling rates, sweep range, noise figure and frequency ranges.

To make the spectrum monitoring system sustainable and permanent, yet able to evolve over time, it must be considered a system-of-systems, where the entire system has an indefinite lifetime, while the individual monitoring elements that comprise it have finite lifetimes. For example, if a receiver in the system becomes outdated it may be replaced by a current one without requiring the replacement of the whole system.

3.2.7 Local and central databases

Distributed site databases can contain detailed information on smaller geographic areas and can be updated quicker to support spectrum sharing**.** A centralized database can leverage monitoring results and data from various sites to drive analytics. It is not a question of one or the other but leveraging both, depending on the situational need. The ability to merge spectrum monitoring data will enable fine-grained analysis of spectrum occupancy at the national, regional, urban and local level.

4 Big data spectrum monitoring

Spectrum monitoring networks need the ability to collect and integrate large amounts of data. The influx of data (volume, variety, and velocity) will strain traditional operational procedures. An information management approach augmented by AI will result in continuous cataloguing of the spectrum on a real-time basis. This data supports analytics, real-time dashboards and mapping tools to enable visualization of current and historical spectrum usage, showing coverage areas, interference areas and unused frequencies:

- Applying AI techniques to learn the spectrum environment;
- Information on unused and under-used frequency bands;
- Interference reports with geolocation and signal recognition;
- Predictive analytics and data mining to assess KPIs on spectrum utilization, spectrum efficiency, spectrum allocation, spectrum users, available spectrum, and interferences;
- Spectrum/frequency allocation alternatives and improvements based on user requirements.

4.1 Big data spectrum monitoring benefits

1 On-demand survey of spectrum utilization and classification (primary, secondary, licensed, unlicensed, static or dynamic).

2 Quantification of the available spectrum by spectrum bands and location (global, regional and local).

3 Coverage analysis for QoS and QoE assessment.

4 Identification, geolocation and cataloguing of interference sources and signals of interest.

5 Detection of unauthorized transmissions, jamming devices and rogue services.

6 Enable proactive monitoring to enhance total spectrum usage and spectral efficiency by exploiting underutilized spectrum with dynamic allocation where possible.

7 Implement ITU-recommended electromagnetic monitoring, technical verification, Direction Finding (DF), and Geolocation measurements.

8 Ensure a fairly accurate signal detection and DF / geolocation.

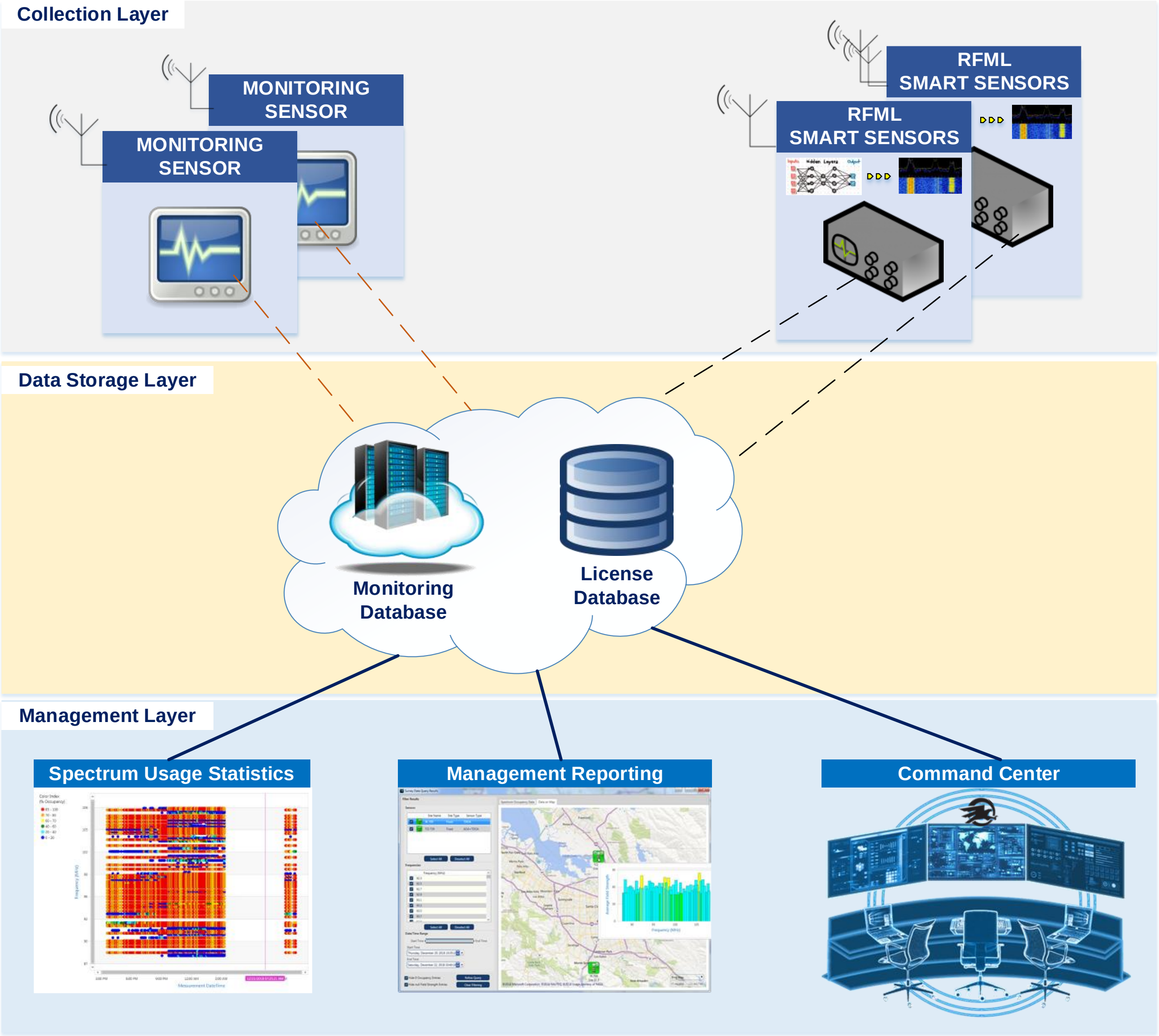
9 Enable future expansion/integration of sensors from different manufacturers.

4.2 Big data spectrum monitoring solution

The solution as shown in Fig. 3 is an example of an intelligence-enabled monitoring network, a central database layer, and an application layer for fine grained monitoring and enforcement of the electromagnetic spectrum covering U/V/S/EHF frequency spectrum.

FIGURE 3

Big data spectrum monitoring solution – a conceptual view

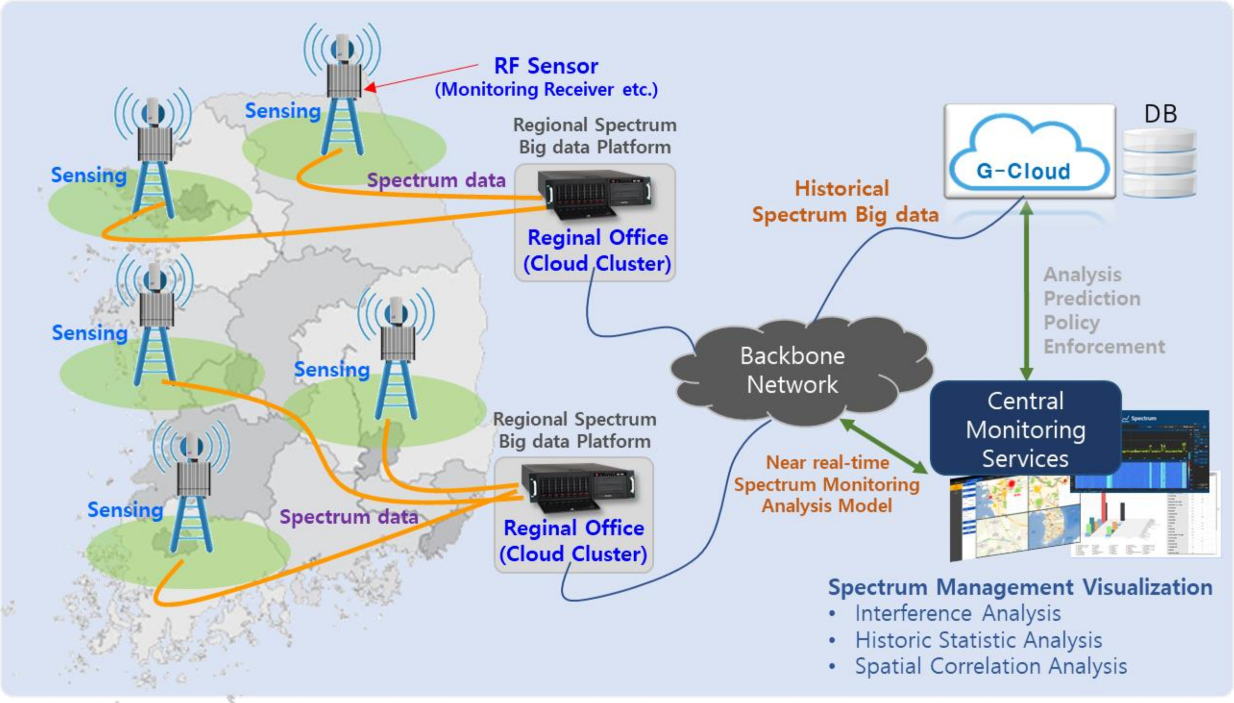


4.3 RF collection layer: example big data spectrum monitoring network

A big data spectrum monitoring network necessitates an overall change in the industry from equipment-oriented systems to data processing-oriented systems. The RF monitoring sensor will evolve from basic spectrum monitoring functions to one equipped with processing intelligence to survey spectrum, catalogue signals, detect interference and capture violations.

FIGURE 4

National big data spectrum monitoring network



The spectrum big data platform illustrated in Fig. 4 includes a network of RF monitoring sensors capable of supporting existing monitoring assets along with new sensors that offer enhanced coverage and intelligent processing. A centralized big data repository collects the spectrum monitoring data, violations, alarms, geolocation data and signal identification to enable users to make informed decisions or allow the system to autonomously make decisions.

4.4 Data storage layer

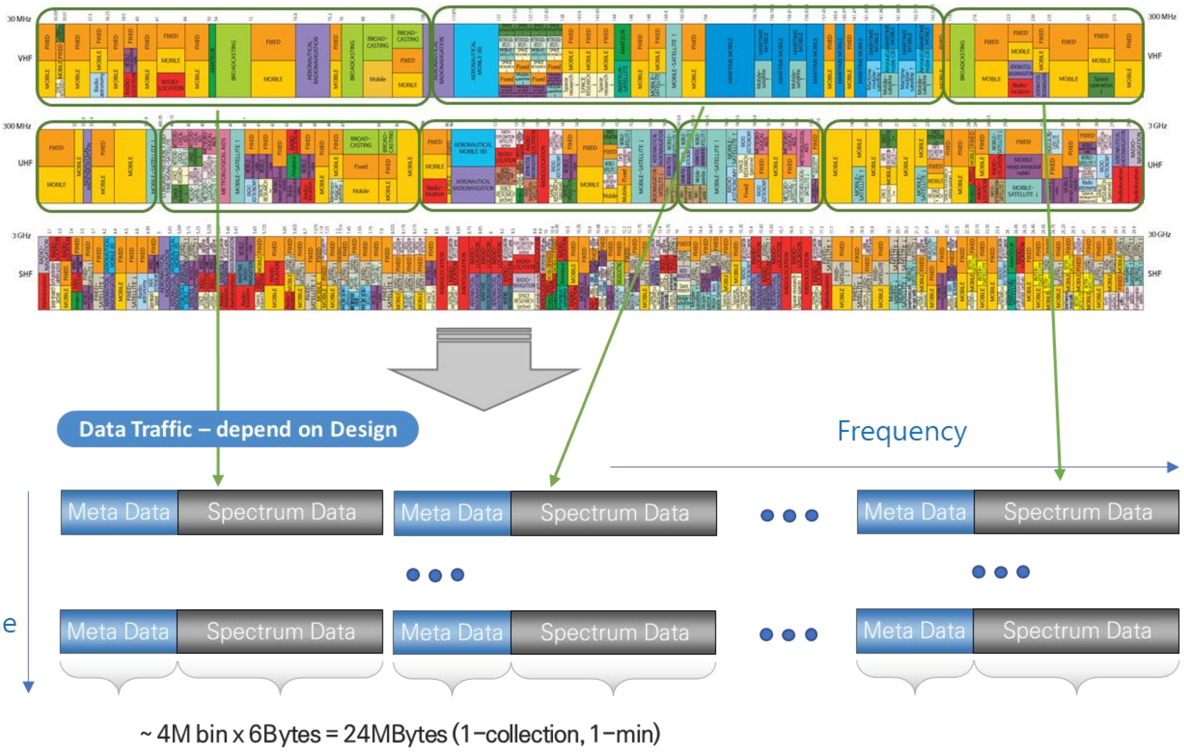
Information in the central database should be stored using industry standard formats and unique identifiers so that the data from various different monitoring systems can be merged into a common database.

It is imperative that a big data spectrum monitoring solution enables the operator to undertake coordinated spectrum measurement campaigns and the ability to task monitoring sensors to perform one or more functions in tandem or in parallel. This enables operators to easily prepare Monitoring Task Plans (MTPs) for multiple stations and provision the same on a regional or local level.

All the information stored in the central database is stored using unique identifiers so that the data from stand-alone and third-party systems can be merged into the database.

FIGURE 5

Data format for spectrum big data gathering

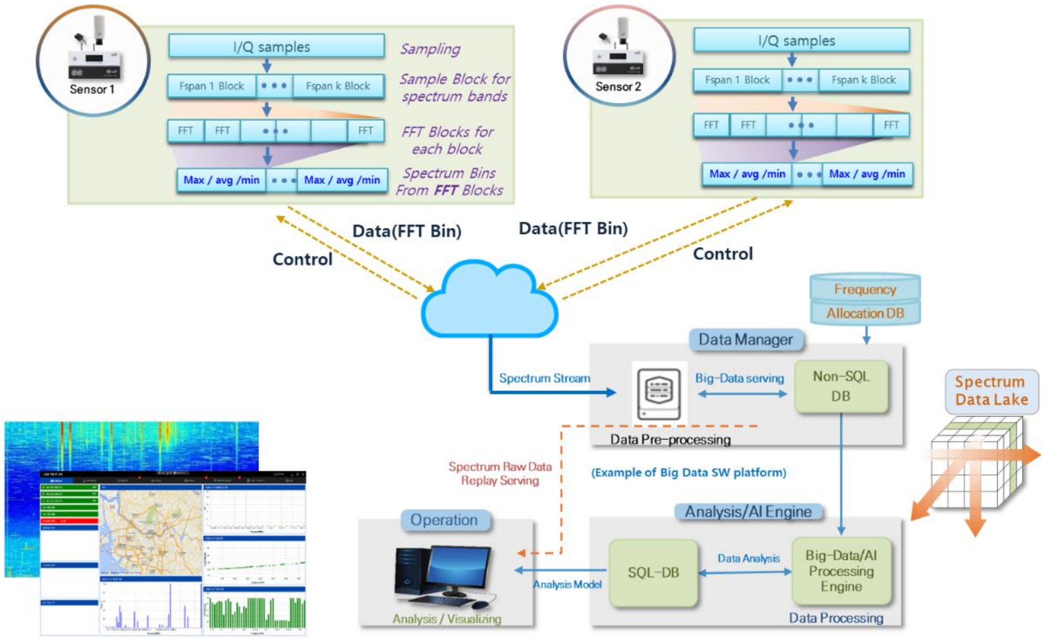


4.4.1 Collection of big data (definition of collected data)

The collection and processing format, illustrated in Fig. 5, consists of both meta data and spectrum data. The collected data is processed into accumulated statistical spectrum data such as maximum, average, and minimum values by processing sequential FFT spectrum data from RF sensors.

FIGURE 6

Example architecture of internet-based spectrum big data gathering process



As described in more detail in Fig. 6, the spectrum collection system is configured to repeatedly process the entire band of the spectrum monitoring receiver (e.g. 20 MHz to 6 GHz) and transmit the statistically processed spectrum data such as maximum, minimum, and average traces during the collection time.

4.4.2 Data pre-processing

The parameters necessary for spectrum monitoring tasks are processed into a pre-processing format for spectrum big data analysis by time series, frequency, and location based on license data (for example) to improve the speed of analysis. In addition, the local processing in each of the RF sensors enables detection of illegal transmissions and interference in near real time.

4.4.3 Data analysis and visualization

Spectral raw data and pre-processed data can be developed for data processing and visualization (as in Fig. 7) by various open source software packages for storage, analysis, and processing.

Spectrum data should be managed for each location and support detailed analysis of monitoring tasks and other activities through post-processing. Longer-term data can be stored using compression methods such as compression algorithms and statistical processing of spectrum raw data.

FIGURE 7

Spectrum monitoring with local processing for illegal transmission and interference detection

X: 974.420455 Y: 2021-03-25 15:41

● max X: 919.8 Y: -99.11
● avg X: 919.8 Y: -105.02
● min X: 919.8 Y: -126.67

☑ max
☑ avg
☑ min

Ch PWR : -62.6dBm
SNR : 16.39dB

2021-03-03 PM 3:03:00

Center Freq : 919. BW : 200kHz

Trace : avg

X: 920.544348 Y: -19.600635

Time	Center Freq(MHz)	BW(kHz)	SNR(dB)
2021-03-25 15:31:00	915	100000	31.44
2021-03-25 15:30:00	915	100000	30.91
2021-03-25 15:29:00	915	100000	32.05
2021-03-25 15:28:00	915	100000	31.11
2021-03-25 15:27:00	915	100000	34.38
			30.65
			34.91
			34.41
			30.12
			32.07
			31.69
			31.28

		(MHz)	Level (dBm)
			-68.5
			-60.8
			-69.9
			-67
			-66.9
			-62.7
			-68.4
2021-03-03 15:06:00	18.2	919.697	-66.8
2021-03-03 15:11:00	15.98	919.698	-62.5
2021-03-03 15:13:00	19.3	919.698	-61.2
2021-03-03 15:14:00	17.51	919.698	-61.7
2021-03-03 15:15:00	14.25	919.7005	-69.7
2021-03-03 15:16:00	12.81	919.7005	-69.9
2021-03-03 15:22:00	11.81	919.6975	-66.9
2021-03-03 15:33:00	13.69	919.7	-70
2021-03-03 15:35:00	10.45	919.698	-65.3
2021-03-03 15:47:00	9.71	919.7	-69.3

Sensing Site
Kyeong-GI Site1

Analysis Freq
Frequency: 915 MHz
BW: 100 MHz

Analysis Period
Start: 2021-03-01 AM 12:00:00
End: 2021-03-06 AM 12:00:00

Manual Analysis
Frequency: 919.7 MHz
BW: 100 kHz
SNR Threshold: 3.0 dB

Manual Filtering
Frequency: 919.7 MHz
BW: 100 kHz
Level Threshold: -70.0 dBm

4.5 Data management layer

System management functions should be easy to use, facilitate collaboration, and enable data integration and analysis.

The need to automate monitoring tasks, group sensors based on functional capabilities and geographic proximity will be needed to enable concurrent collection of data from one or more radio monitoring sensors to gather fine grained data on a national, regional, urban or local level.

Comparative big data spectrum monitoring techniques enable the operators to generate views of spectrum activity across geography and time. The operators undertake coordinated spectrum measurement campaigns with the ability to task monitoring sensors to perform one or more functions in tandem or in parallel. Cooperative measurements are automated by grouping sensors based on functional capabilities and geographic proximity.

The use of cooperative groups of sensors working as teams can also provide real time views of activity and further enable pro-active monitoring and alarms to changes in the environment. These changes may identify unexpected use, overuse or loss of service including interference events.

The use of intuitive GUIs enables the user to:

- Manage cooperative spectrum monitoring teams;
- Manage tasking;
- Query results (data analytics);
- Set the time period for measurements;
- Set the geographical area and location;
- Integrate multiple data sets and databases to facilitate analysis in support of decision making.

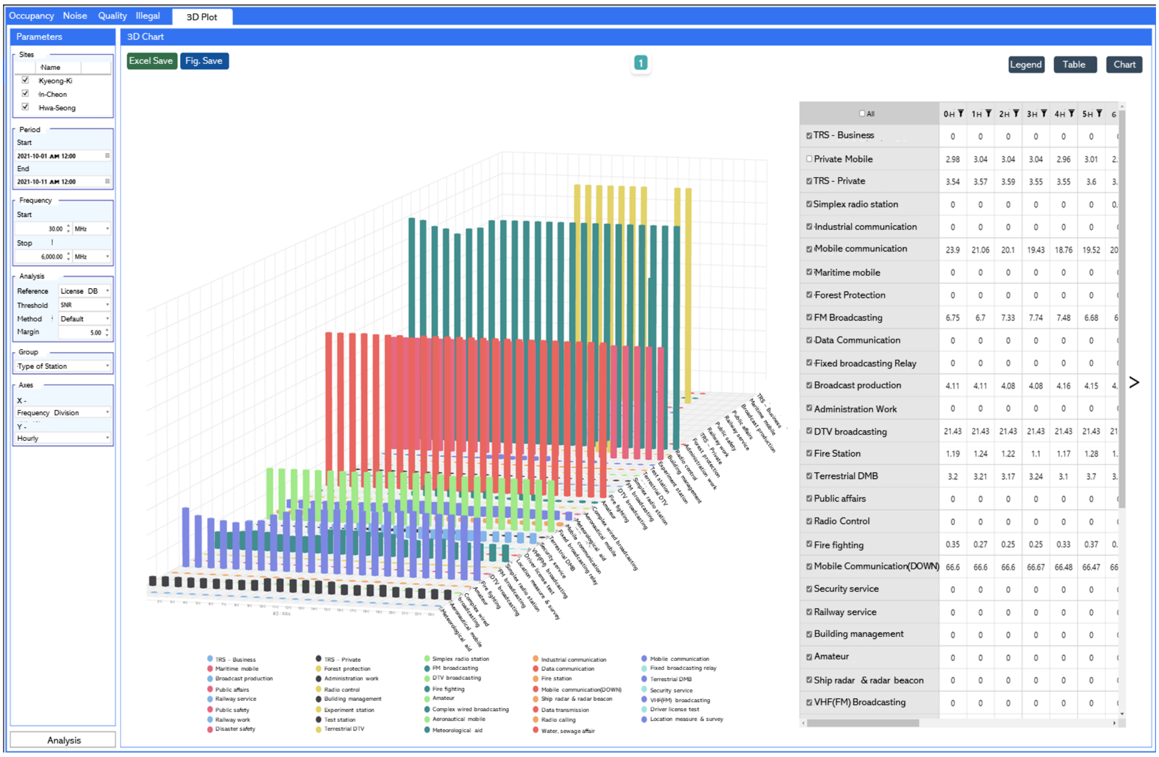
Integration of different systems, through the use of APIs will support commercially available Business Intelligence (BI) / database reporting tools which provide an interface to build dashboards to gain access to a data schema. Published data schema make it possible to create new views and perspectives on the data. This also enables the integration of multiple data sets and associated views.

Through this data schema and the integration of BI tools, users are able to store and provide wide access to the monitoring data. The data can be accessed by multiple clients, each with their own customized view of the data of interest for the intended consumer, organization, or department. With the availability of both spectrum monitoring and geolocation data, dashboards can show the results of the various datasets in very informative and insightful ways. The datasets allow for spectrum occupancy data to be analysed across a band combined with geographical information. This data can be filtered and rearranged using various dashboard tools (see Figs 8 to 11).

Annex 1 to this report provide an example and use case of AI and Big Data, describing the experience of Korea (Republic of) with the use of AI and big data for spectrum monitoring, including a typical intuitive GUI.

FIGURE 8

Frequency occupancy based on hourly intervals



Sites

	Name
☑	Kyeong-Ki
☑	In-Cheon
☑	Hwa-Seong

Period
Start: 2021-10-01 AM 12:00
End: 2021-10-11 AM 12:00

Frequency
Start: 30.00 MHz
Stop: 6,000.00 MHz

Analysis
Reference: License DB
Threshold: SNR
Method: Default
Margin: 5.00

Group: Type of Station

Axes
X - Frequency Division
Y - Hourly

☐ All	0H	1H	2H	3H	4H	5H
☑ TRS - Business	0	0	0	0	0	0
☐ Private Mobile	2.98	3.04	3.04	3.04	2.96	3.01
☑ TRS - Private	3.54	3.57	3.59	3.55	3.55	3.6
☑ Simplex radio station	0	0	0	0	0	0
☑ Industrial communication	0	0	0	0	0	0
☑ Mobile communication	23.9	21.06	20.1	19.43	18.76	19.52
☑ Maritime mobile	0	0	0	0	0	0
☑ Forest Protection	0	0	0	0	0	0
☑ FM Broadcasting	6.75	6.7	7.33	7.74	7.48	6.68
☑ Data Communication	0	0	0	0	0	0
☑ Fixed broadcasting Relay	0	0	0	0	0	0
☑ Broadcast production	4.11	4.11	4.08	4.08	4.16	4.15
☑ Administration Work	0	0	0	0	0	0
☑ DTV broadcasting	21.43	21.43	21.43	21.43	21.43	21.43
☑ Fire Station	1.19	1.24	1.22	1.1	1.17	1.28
☑ Terrestrial DMB	3.2	3.21	3.17	3.24	3.1	3.7
☑ Public affairs	0	0	0	0	0	0
☑ Radio Control	0	0	0	0	0	0
☑ Fire fighting	0.35	0.27	0.25	0.25	0.33	0.37
☑ Mobile Communication(DOWN)	66.6	66.6	66.6	66.67	66.48	66.47
☑ Security service	0	0	0	0	0	0
☑ Railway service	0	0	0	0	0	0
☑ Building management	0	0	0	0	0	0
☑ Amateur	0	0	0	0	0	0
☑ Ship radar & radar beacon	0	0	0	0	0	0
☑ VHF(FM) Broadcasting	0	0	0	0	0	0

FIGURE 9

3D visualization of spectrum occupancy over a frequency range and time

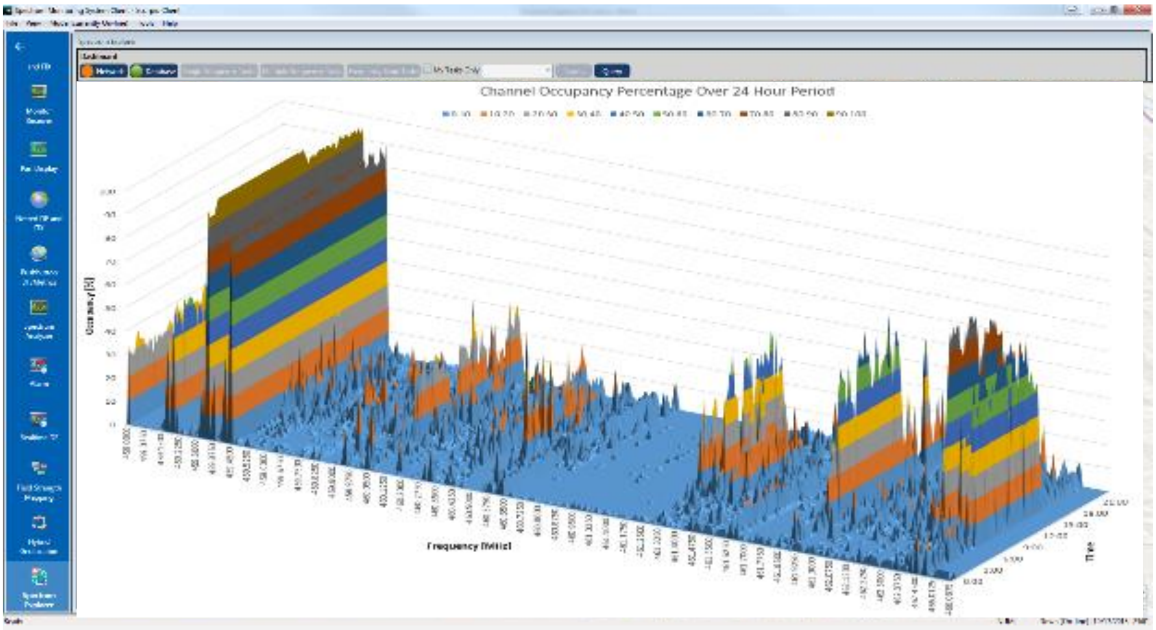
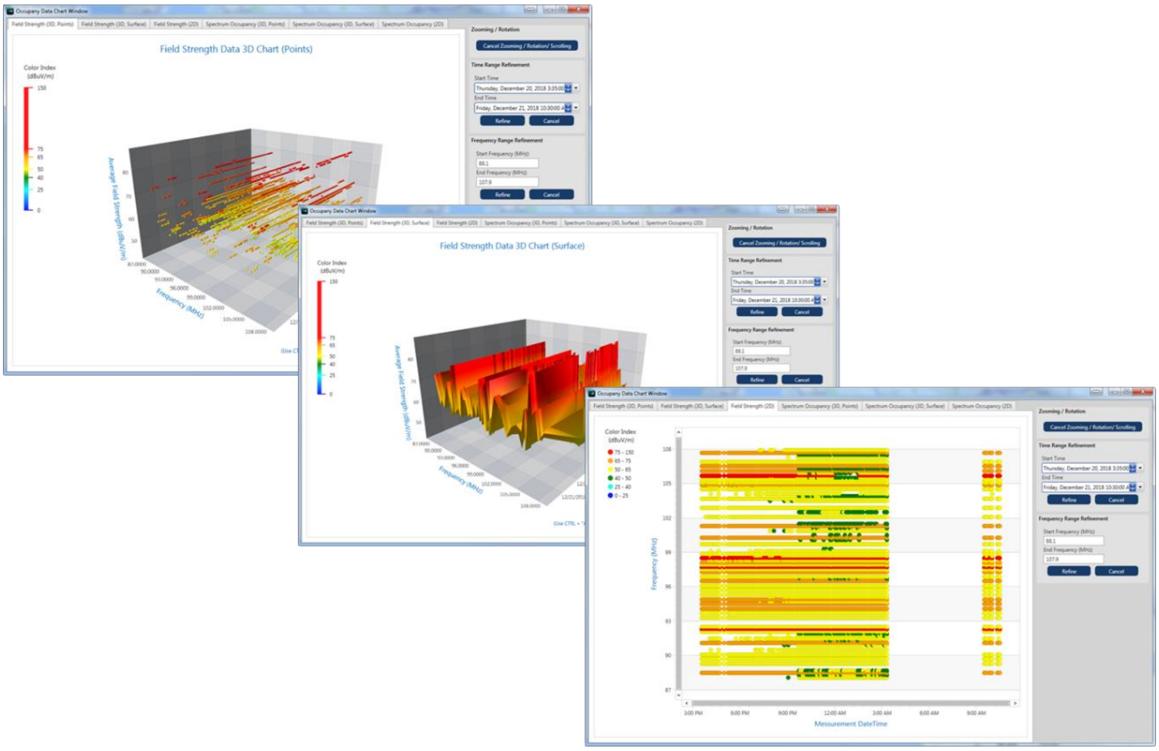


FIGURE 10

Built-in query for evaluating spectrum occupancy in 2D and 3D formats



Start Time: Thursday, December 20, 2018 3:35:00

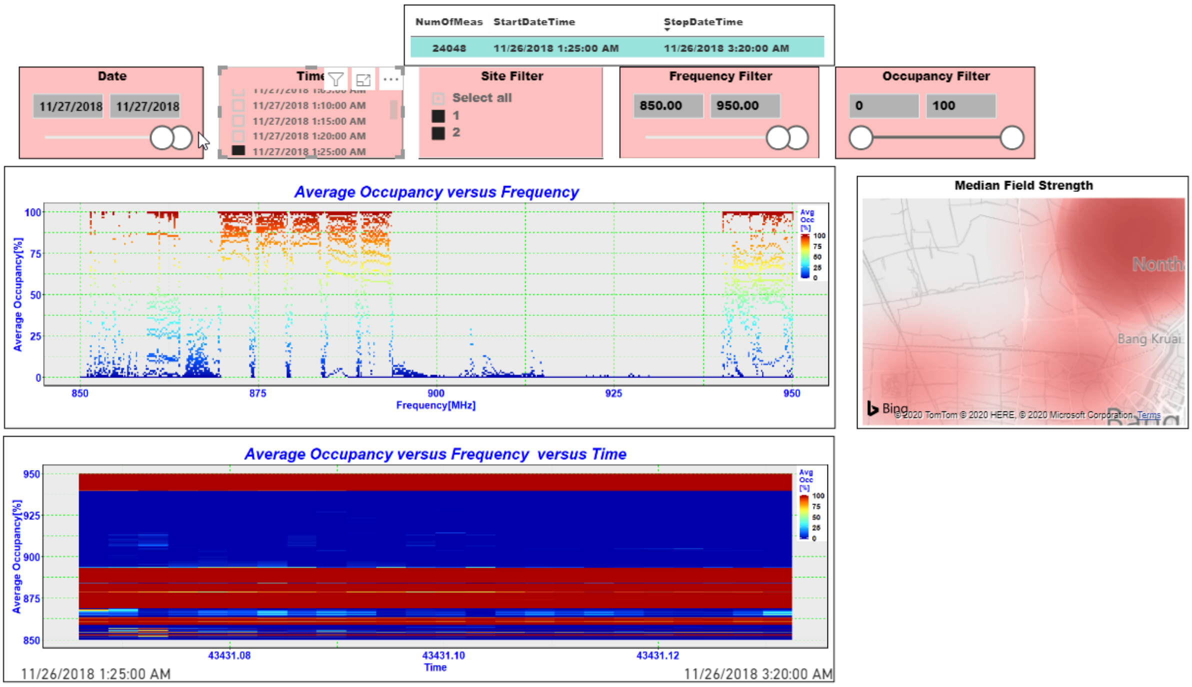
End Time: Friday, December 21, 2018 10:30:00 A

Start Frequency (MHz): 88.1

End Frequency (MHz): 107.9

FIGURE 11

Dashboards for spectrum assessment by frequency, time and location



NumOfMeas	StartDateTime	StopDateTime
24048	11/26/2018 1:25:00 AM	11/26/2018 3:20:00 AM

Date
11/27/2018 11/27/2018

Time
11/27/2018 1:10:00 AM
11/27/2018 1:15:00 AM
11/27/2018 1:20:00 AM
11/27/2018 1:25:00 AM

Site Filter
☐ Select all
■ 1
■ 2

Frequency Filter
850.00 950.00

Occupancy Filter
0 100

11/26/2018 1:25:00 AM
11/26/2018 3:20:00 AM

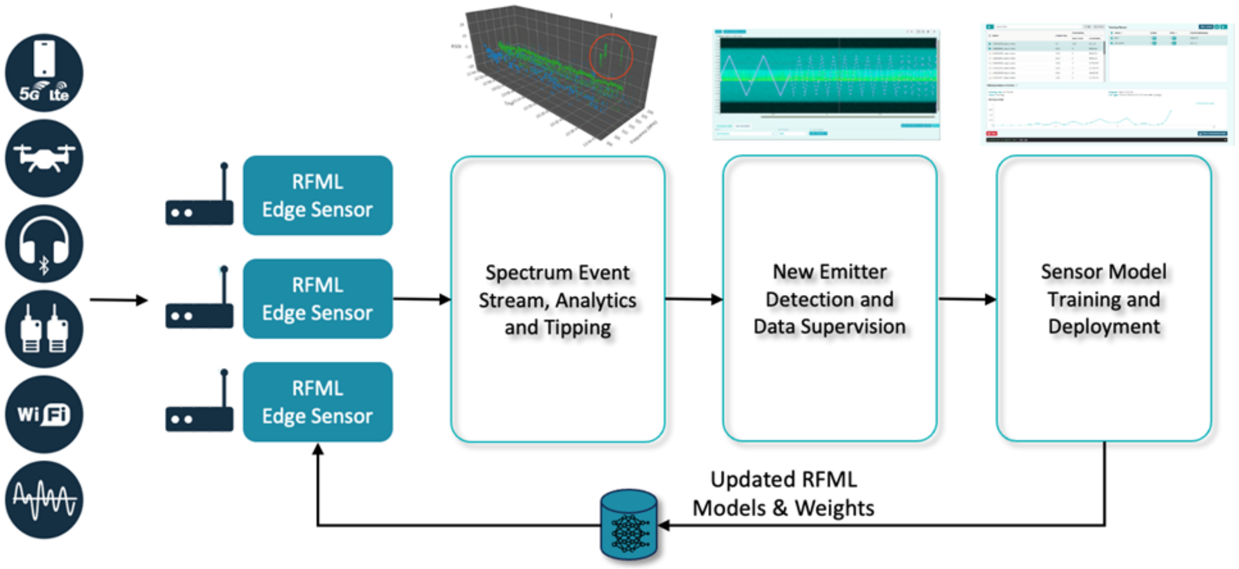
5 Realtime data-driven spectrum awareness using RFML

With advances in wireless communications and wireless devices, and increasing diversity of new and adaptive wireless systems, it is important to not only detect and classify existing protocols, but to also identify, learn and adapt to future protocols and wireless systems as they are deployed – including those in mobile/cellular networks, IoT networks, autonomous systems, and other emerging technologies.

RFML signal processing can use neural networks to learn RF protocol and emitter characteristics in order to rapidly detect and classify their type [4]. As new wireless systems, services, and protocols are introduced, RFML based spectrum sensing systems can identify new and anomalous types of emissions, help to label, and annotate them in datasets, and train to identify them when seen in the future. Further, microprocessor hardware has increasingly become optimized for cost and energy efficient neural network inference and can be deployed to the edge of the monitoring network. Deploying RFML at the edge provides near real-time sensing, with the ability to efficiently detect and identify numerous spectrum events, providing an efficient and effective way to audit and analyse activity over broad frequency bands and long-time intervals. This can be done without the need for storage or transmission of raw I/Q time-series data files. A data-driven workflow describing the interactions of edge inference, event analytics, data curation, model retraining and deployment for new signals and emissions is shown in Fig. 12.

FIGURE 12

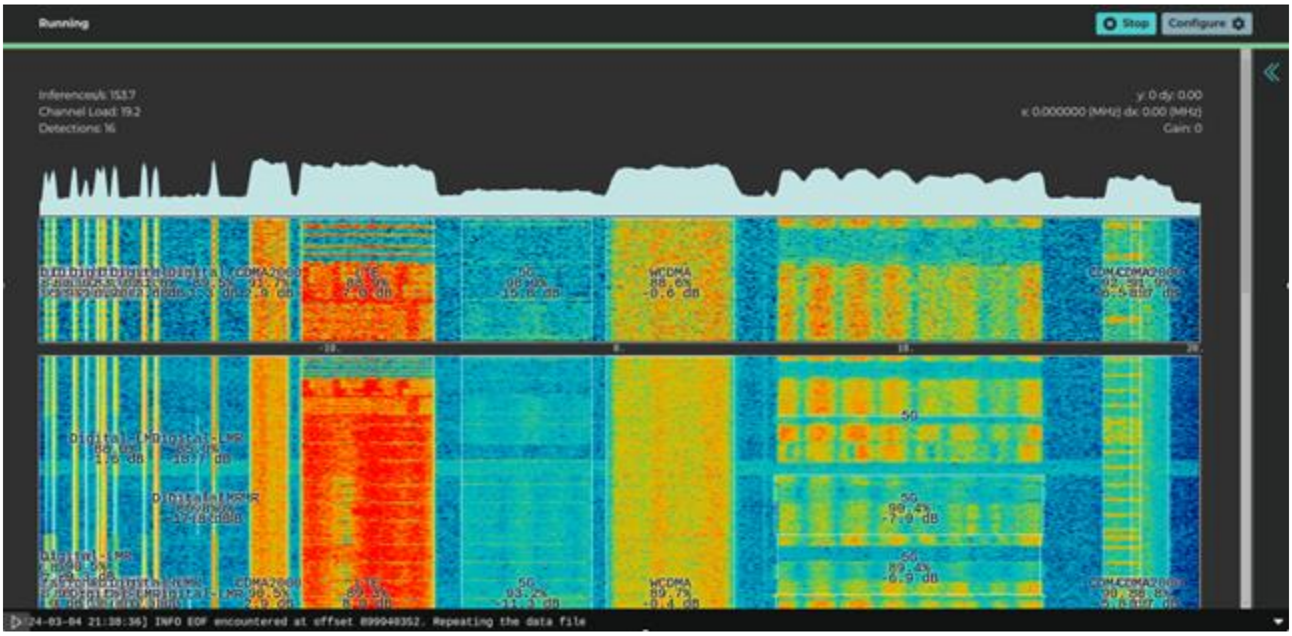
Data-driven spectrum sensing workflow



Leveraging AI through RFML-based spectrum sensing enables efficient and scalable signal detection, identification, analytics, anomaly, and trend detection with near-real-time insight into signal activity and band occupancy. It provides a smart view into signals activity across a wide range of bands and technologies and event streams of metadata describing signal types, power levels, properties and behaviours. This can be leveraged by automation at scale or visualized by users as shown in the spectrogram illustrating the detection, identification, and annotation of RF emissions in a wide band. (See Fig. 13).

FIGURE 13

Sample graphical output of RFML-based sensor showing signal classification and confidence level



```
24-03-04 21:38:36] INFO EOF encountered at offset 899940352. Repeating the data file
```

AI driven spectrum monitoring has the potential to offer fully autonomous 24/7 signal detection and recognition and analytics from multiple monitoring sites to support:

- Confirmation that licensed protocols and systems are being used;
- Confirmation that spectrum allocations are being actively used;
- Interference detection, identification, and feedback;
- Available spectrum assessment and feedback for optimization;
- Enablement of air-interface digital twin information;
- Rapid learning of new types of emitters and protocols;
- Automated analytics and monitoring of wide range of trends in spectrum access.

6 Summary

There have been many advancements on the application of ML and AI to spectrum monitoring and there is a large volume of technical papers outlining the use, the algorithms, and the limitations of the technology, as well as commercially available software solutions. These developments will likely lead to more autonomous spectrum monitoring systems that are not only based on classic automation, but also on intelligent, learning capabilities able to collect and fuse data from multiple sources (e.g. spectral data, environment, traffic, events, etc.) and make rapid and good decisions at scale.

Dense distributed monitoring networks produce high quality and real-time occupancy data that makes the discovery of available spectrum across the covered area possible. It also could support accurate predictions for autonomous allocation of frequencies for specific wireless applications like IoT [2]. In some cases, spectrum sensing capabilities could be co-located and co-deployed with network infrastructure such as within base station and radio unit hardware, and could where appropriate, feasible and applicable coordinate with network elements.

Intelligent distributed spectrum monitoring systems hold the potential to improve the efficiency of regulatory personnel by reducing the effort to obtain the monitoring data required to manage the spectrum and allowing insight into spectrum activity trends across many locations, bands, types of emitters, and timescales which were previously unmanageable through other means.

In order to build a large-scale spectrum monitoring and automated spectrum allocation system, it is necessary to integrate spectrum sensing data through common rules and standardization of measurements. It should be possible to measure the related RF frequencies for each site using different spectrum sensors, and to distribute them in a standardized integrated format in order to obtain the trust of spectrum big data analysis. Use of standard protocols for RF datasets can result in interoperability and efficient deployment, integration, and automation of such systems on a broad scale.

7 References

[1] Wong, Lauren J., IV William H.Clark, Bryse Flowers, R. Michael Buehrer, Alan J. Michaels and William C. Headley. "The RFML Ecosystem: A Look at the Unique Challenges of Applying Deep Learning to Radio Frequency Applications." arXiv: Signal Processing (2020): n. pag. Online: https://arxiv.org/pdf/2010.00432

[2] Recommendation ITU-R SM.1046 – Definition of spectrum use and efficiency of a radio system

[3] Baltiiski, P., Iliev, I., Kehaiov, B. et al. Long-Term Spectrum Monitoring with Big Data Analysis and Machine Learning for Cloud-Based Radio Access Networks. Wireless Pers Commun 87, 815–835 (2016). https://doi.org/10.1007/s11277-015-2631-8

[4] T. J. O'Shea, T. Roy and T. C. Clancy, "Over-the-Air Deep Learning Based Radio Signal Classification," in IEEE Journal of Selected Topics in Signal Processing, vol. 12, no. 1, pp. 168-179, Feb. 2018, doi: 10.1109/JSTSP.2018.2797022

Annex 1

Solution for data-driven AI and big data spectrum monitoring in Korea (Republic of)

Korea is performing spectrum monitoring through big data analysis, using a dedicated software program which can realize the data-driven big data analysis. The spectrum is monitored by three monitoring sites in Korea as a pilot system. Each cloud cluster server uses sensors covering 30 MHz to 7.5 GHz to collect spectrum data for building a spectrum database.

The government of Korea plans to expand the data-driven big data spectrum-monitoring system to the whole country in the future. There is a clear direction that spectrum management should include big data and AI (Artificial Intelligence) technology and have a global perspective.

FIGURE A1-1

Spectrum big data analysis workflow

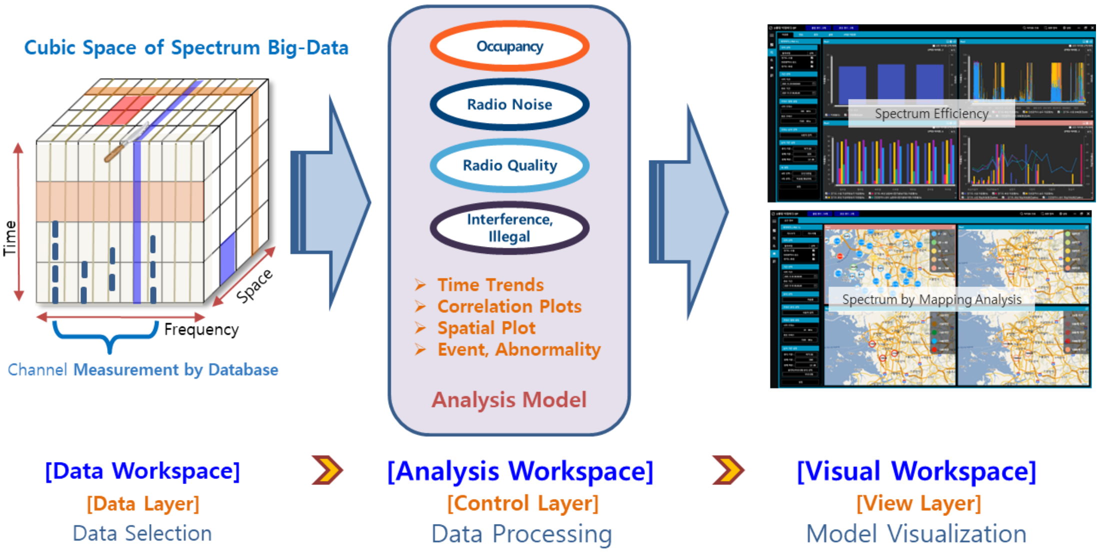


Figure A1-1 illustrates the spectrum big data analysis workflow including data workspace, analysis workspace and visual workspace.

The RF sensors autonomously collect designated spectrum data considering the national frequency allocation table and radio license database. The system produces spectrum data which calculates through an FFT process the statistical values including maximum, minimum, and average. The generated spectrum data is periodically transferred to a cloud big data server. The cloud server saves two forms of the spectrum data including raw data and pre-processing data. When spectrum data is saved at the cloud server, it is stored as meta data which can distinguish each parameter. The meta data consists of several variables as listed in the table below. This spectrum data can be accessed at any regional office and central office. Administrators can easily access the spectrum data due to its organized meta data form.

TABLE A1-1

Meta Data Format

Class	Category	Type
Sensor & General Information	Sensor type	String
	Sensor ID	Integer
	Operator’s management information	String
	Antenna model	String
	Spectrum measurement time	Float
	Spectrum form	Integer
	Sensor temperature	Float
	Sensor RF status	Integer array
	Measurement time zone	String

TABLE A1-1 (*end*)

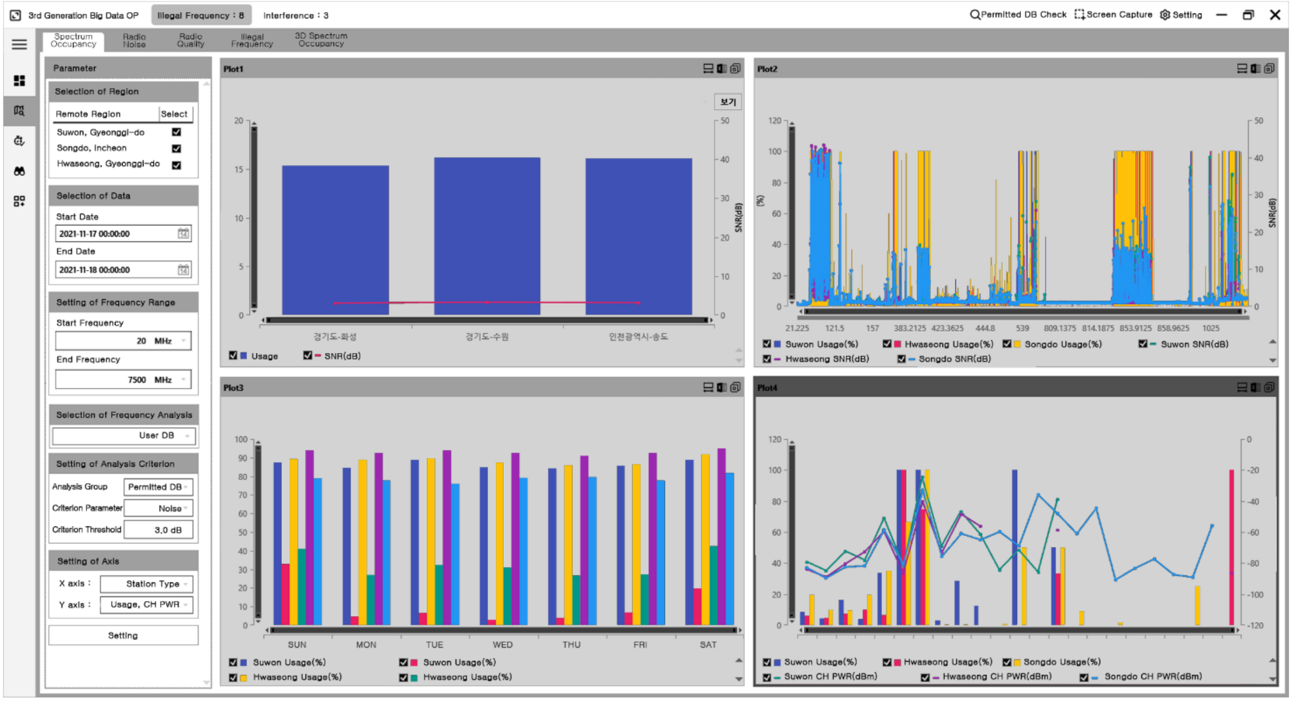
Class	**Category**	**Type**
Measurement Information	Measurement date and time	String
	Measurement longitude	Float
	Measurement latitude	Float
	Measurement height	Float
	GPS lock status	Integer
	Measurement frequency	Double
	Measurement bandwidth	Float
	Reference level	Integer
	Number of spectrum bins	Integer
	Interval of spectrum bins	Float
	Signal existence flag	Integer array
Signal Information	Average noise level	Float array

The raw spectrum data serves as a near real time streaming spectrum form for interference and illegal frequency detection using AI or other algorithms. The pre-processing data is used for fast big data analysis of spectrum monitoring tasks such as occupancy, radio noise and radio quality.

Figure A1-2 shows display examples for occupancy, radio noise, radio quality, interference and illegal frequency analysis.

FIGURE A1-2

Major analysis factors of the data-driven big data spectrum monitoring solution



3rd Generation Big Data OP
Illegal Frequency : 8
Interference : 3

Parameter

Selection of Region

Remote Region	Select
Suwon, Gyeonggi-do	☑
Songdo, Incheon	☑
Hwaseong, Gyeonggi-do	☑

Selection of Data

Start Date: 2021-11-17 00:00:00

End Date: 2021-11-18 00:00:00

Setting of Frequency Range

Start Frequency: 20 MHz

End Frequency: 7500 MHz

Selection of Frequency Analysis: User DB

Setting of Analysis Criterion

Analysis Group: Permitted DB

Criterion Parameter: Noise

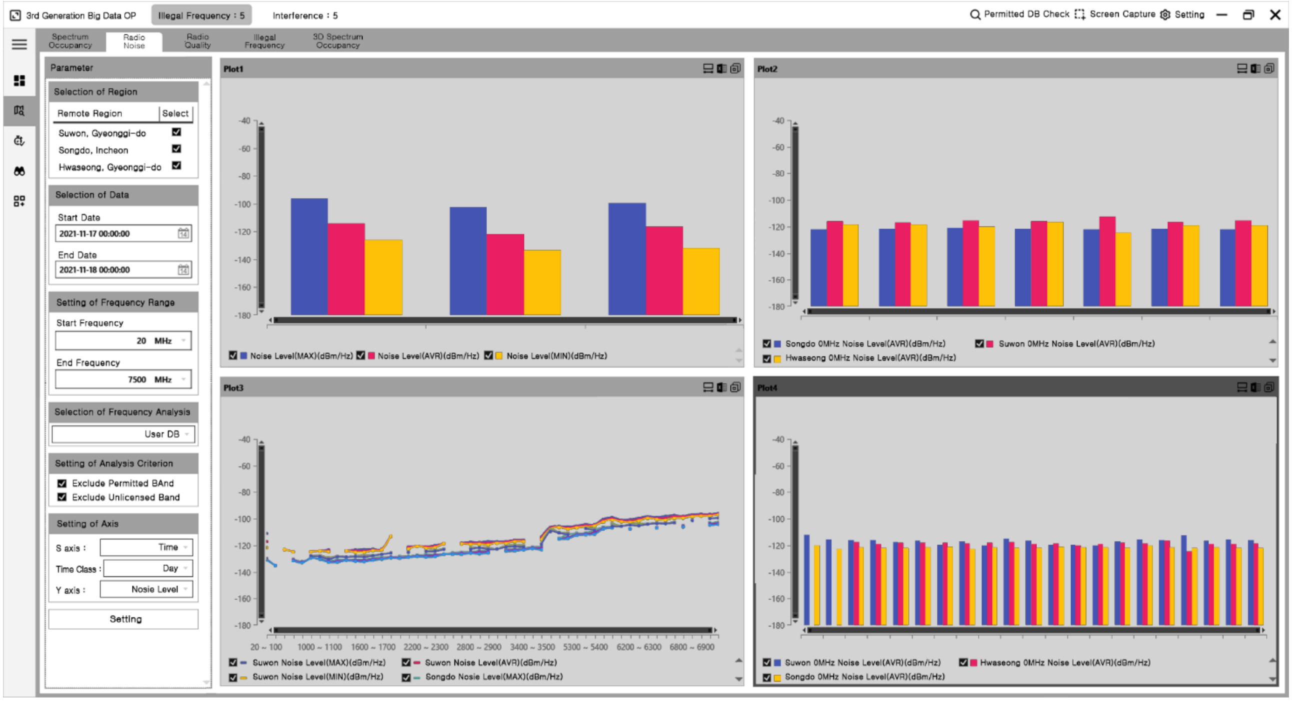
Criterion Threshold: 3,0 dB

Setting of Axis

X axis : Station Type

Y axis : Usage, CH PWR

a. Spectrum occupancy



Parameter

Selection of Region

Remote Region	Select
Suwon, Gyeonggi-do	☑
Songdo, Incheon	☑
Hwaseong, Gyeonggi-do	☑

Selection of Data
Start Date 2021-11-17 00:00:00
End Date 2021-11-18 00:00:00

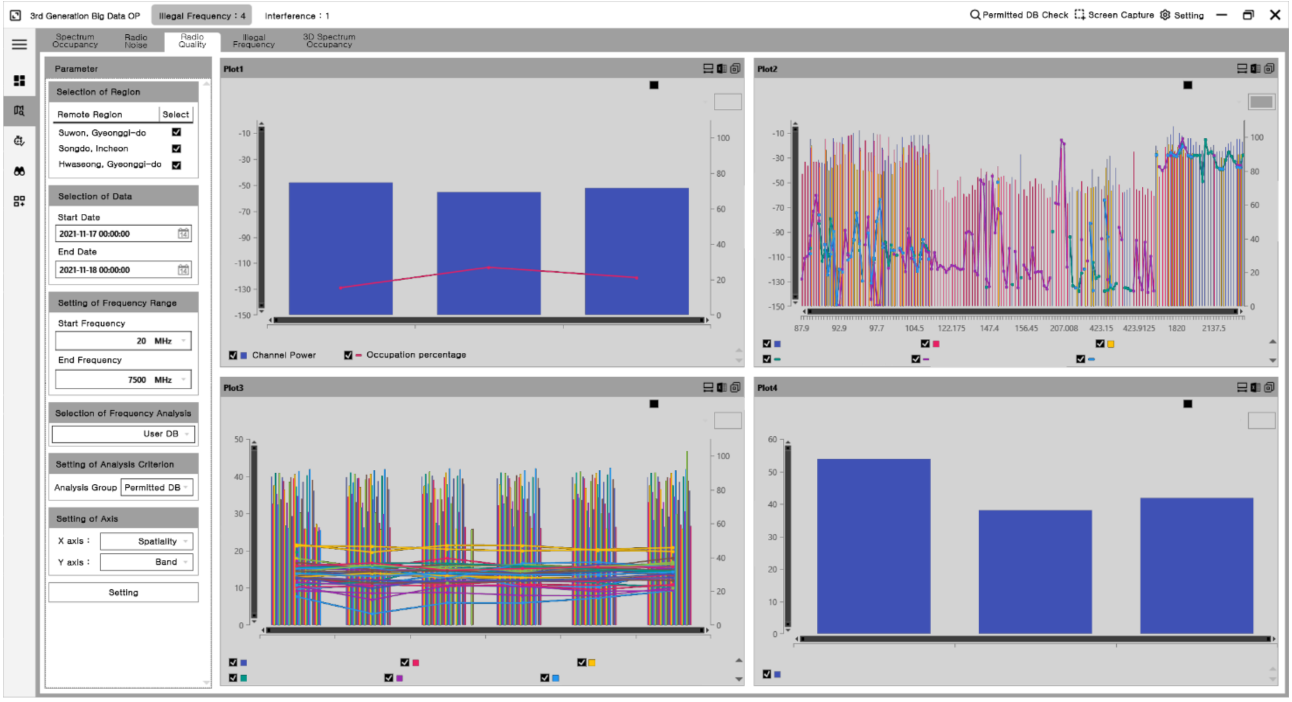
Setting of Frequency Range
Start Frequency 20 MHz
End Frequency 7500 MHz

Selection of Frequency Analysis
User DB

Setting of Analysis Criterion
☑ Exclude Permitted BAnd
☑ Exclude Unlicensed Band

Setting of Axis
S axis : Time
Time Class : Day
Y axis : Nosie Level

b. Radio noise



Parameter

Selection of Region

Remote Region	Select
Suwon, Gyeonggi-do	☑
Songdo, Incheon	☑
Hwaseong, Gyeonggi-do	☑

Selection of Data
Start Date 2021-11-17 00:00:00
End Date 2021-11-18 00:00:00

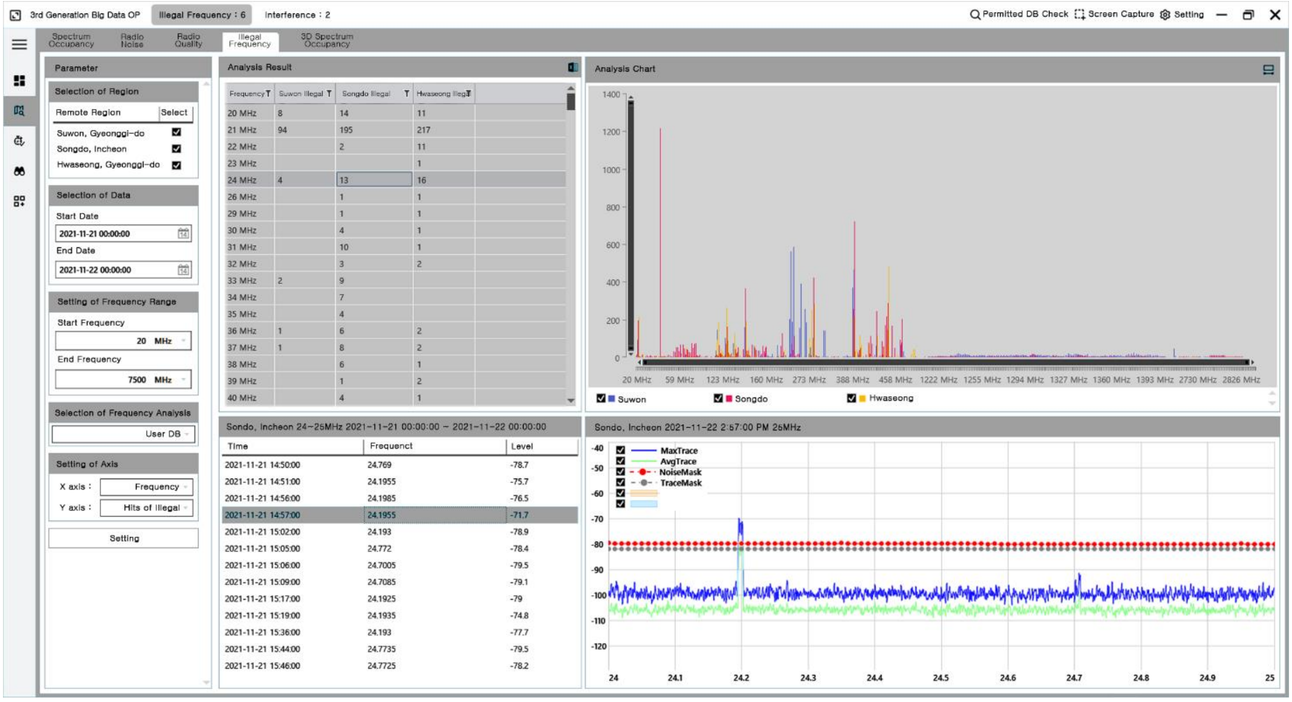
Setting of Frequency Range
Start Frequency 20 MHz
End Frequency 7500 MHz

Selection of Frequency Analysis
User DB

Setting of Analysis Criterion
Analysis Group Permitted DB

Setting of Axis
X axis : Spatiality
Y axis : Band

c. Radio quality



Selection of Region

Remote Region	Select
Suwon, Gyeonggi-do	☑
Songdo, Incheon	☑
Hwaseong, Gyeonggi-do	☑

Selection of Data

Start Date: 2021-11-21 00:00:00

End Date: 2021-11-22 00:00:00

Setting of Frequency Range

Start Frequency: 20 MHz

End Frequency: 7500 MHz

Selection of Frequency Analysis

User DB

Setting of Axis

X axis : Frequency

Y axis : Hits of Illegal

Setting

Frequency	Suwon Illegal	Songdo Illegal	Hwaseong Illegal
20 MHz	8	14	11
21 MHz	94	195	217
22 MHz		2	11
23 MHz			1
24 MHz	4	13	16
26 MHz		1	1
29 MHz		1	1
30 MHz		4	1
31 MHz		10	1
32 MHz		3	2
33 MHz	2	9	
34 MHz		7	
35 MHz		4	
36 MHz	1	6	2
37 MHz	1	8	2
38 MHz		6	1
39 MHz		1	2
40 MHz		4	1

Sondo, Incheon 24~25MHz 2021-11-21 00:00:00 ~ 2021-11-22 00:00:00

Time	Frequenct	Level
2021-11-21 14:50:00	24.769	-78.7
2021-11-21 14:51:00	24.1955	-75.7
2021-11-21 14:56:00	24.1985	-76.5
2021-11-21 14:57:00	24.1955	-71.7
2021-11-21 15:02:00	24.193	-78.9
2021-11-21 15:05:00	24.772	-78.4
2021-11-21 15:06:00	24.7005	-79.5
2021-11-21 15:09:00	24.7085	-79.1
2021-11-21 15:17:00	24.1925	-79
2021-11-21 15:19:00	24.1935	-74.8
2021-11-21 15:36:00	24.193	-77.7
2021-11-21 15:44:00	24.7735	-79.5
2021-11-21 15:46:00	24.7725	-78.2

d. Illegal frequency

A spectrum occupancy screen such as illustrated in Fig. A1-2 is an indicator of how efficiently each allocated spectrum is utilized. The spectrum efficiency plays a significant role in spectrum retrieval and reallocation (spectrum refarming) because the spectrum is a limited resource.

Radio noise environment is another important factor. During spectrum monitoring a combination of the spectrum of interest and the noise is measured. The radio noise refers to random fluctuations of the instantaneous spectrum because of artificial and natural noise.

Radio quality is an additional important factor. An analysis result of radio quality indicates whether or not a signal complies with licensed radio quality parameters such as SNR, channel power, occupancy bandwidth, etc.

A remaining important monitoring factor within the dedicated software program is an autonomous illegal frequency search. This is important since an illegal frequency can influence and disrupt the legal frequency authorized by the administration.

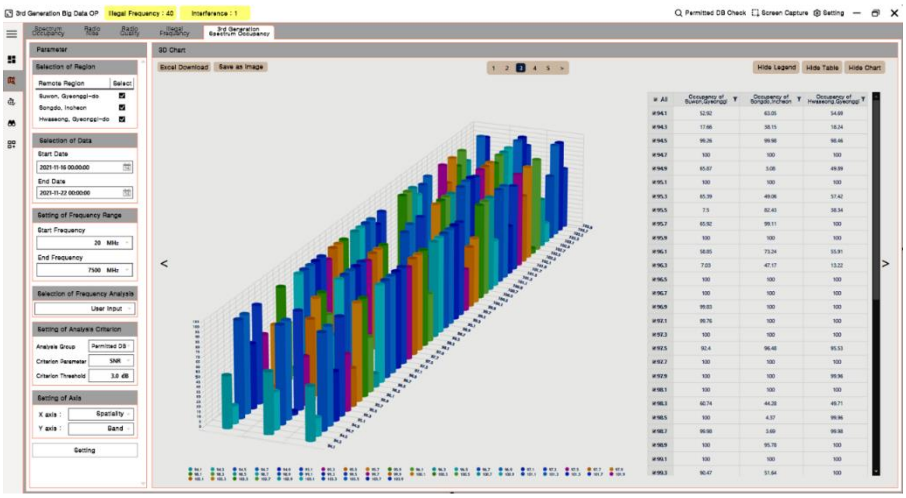
Illegal and interference frequency analysis shows statistical results of time series occurrences and historical data for each frequency band or space. An illegal or interference frequency is detected by using streaming spectrum data from the automatically updated licensed database.

The dedicated software program can analyse the spectrum three-dimensionally, with x, y and z coordinates as illustrated in Fig. A1-3.

Figure A1-3 illustrates spectrum occupancy as measured at three monitoring stations for each axis as x: selected monitoring station(space), y: frequency channel assigned(frequency) and z :spectrum occupancy for assigned period for analysis condition(time).

FIGURE A1-3

Three-dimensional spectrum monitoring analysis



Selection of Region

Remote Region	Select
Suwon, Gyeonggi-do	☑
Songdo, Incheon	☑
Hwaseong, Gyeonggi-do	☑

Selection of Data

Start Date: 2021-11-16 00:00:00

End Date: 2021-11-22 00:00:00

Setting of Frequency Range

Start Frequency: 20 MHz

End Frequency: 7500 MHz

Selection of Frequency Analysis: User Input

Setting of Analysis Criterion

Analysis Group: Permitted DB

Criterion Parameter: SNR

Criterion Threshold: 3.0 dB

Setting of Axis

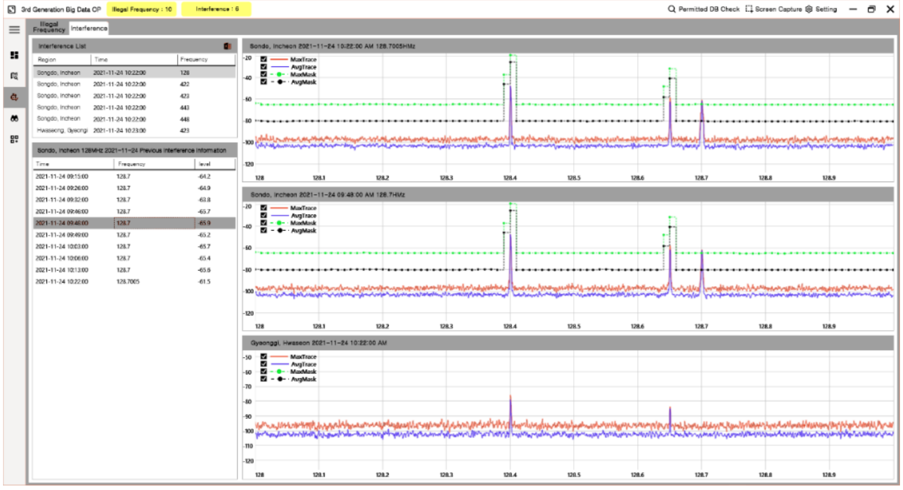
X axis : Spatiality

Y axis : Band

# All	Occupancy of Suwon,Gyeonggi	Occupancy of Songdo,Incheon	Occupancy of Hwaseong,Gyeonggi
#94.1	52.92	63.05	54.69
#94.3	17.66	58.15	18.24
#94.5	99.26	99.98	98.46
#94.7	100	100	100
#94.9	65.87	3.08	49.89
#95.1	100	100	100
#95.3	65.39	49.06	57.42
#95.5	7.5	82.43	38.34
#95.7	65.92	99.11	100
#95.9	100	100	100
#96.1	58.85	73.24	55.91
#96.3	7.03	47.17	13.22
#96.5	100	100	100
#96.7	100	100	100
#96.9	99.83	100	100
#97.1	99.76	100	100
#97.3	100	100	100
#97.5	92.4	96.48	95.53
#97.7	100	100	100
#97.9	100	100	99.96
#98.1	100	100	100
#98.3	60.74	44.28	49.71
#98.5	100	4.37	99.96
#98.7	99.98	3.69	99.98
#98.9	100	95.78	100
#99.1	100	100	100
#99.3	90.47	51.64	100

FIGURE A1-4

Streaming data analysis



Interference List

Region	Time	Frequency
Songdo, Incheon	2021-11-24 10:22:00	128
Songdo, Incheon	2021-11-24 10:22:00	422
Songdo, Incheon	2021-11-24 10:22:00	423
Songdo, Incheon	2021-11-24 10:22:00	443
Songdo, Incheon	2021-11-24 10:22:00	448
Hwaseong, Gyeongi	2021-11-24 10:23:00	423

Sondo, Incheon 128MHz 2021-11-24 Previous Interference Information

Time	Frequency	level
2021-11-24 09:15:00	128.7	-64.2
2021-11-24 09:26:00	128.7	-64.9
2021-11-24 09:32:00	128.7	-63.8
2021-11-24 09:46:00	128.7	-65.7
2021-11-24 09:48:00	128.7	-65.9
2021-11-24 09:49:00	128.7	-65.2
2021-11-24 10:03:00	128.7	-65.7
2021-11-24 10:06:00	128.7	-65.4
2021-11-24 10:13:00	128.7	-65.8
2021-11-24 10:22:00	128.7005	-61.5

The software monitors the illegal frequency and interference spectrum in real time. The streaming data analysis can track previous records which detected illegal or suspicious spectrum.

This analysis technique is illustrated in Fig. A1-4. The suspicious spectrum can be observed as a sudden change in the streaming data flow. Figure A1-4 illustrates the detection of an illegal frequency (upper trace) and the analysis results of a historically similar illegal spectrum occurrence at previous times (lower trace). This analysis can be selected to be compared with a particular region at the current time to roughly estimate location of transmitter.

FIGURE A1-5

Space data analysis by big data spectrum monitoring

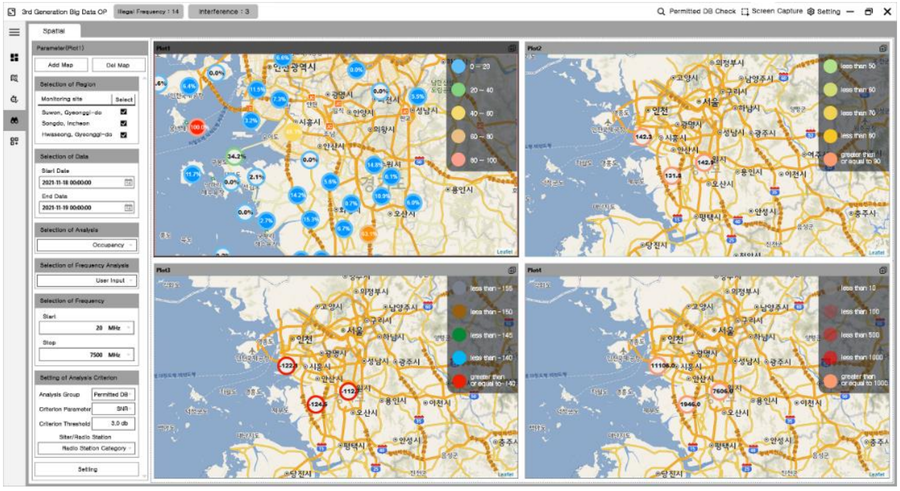
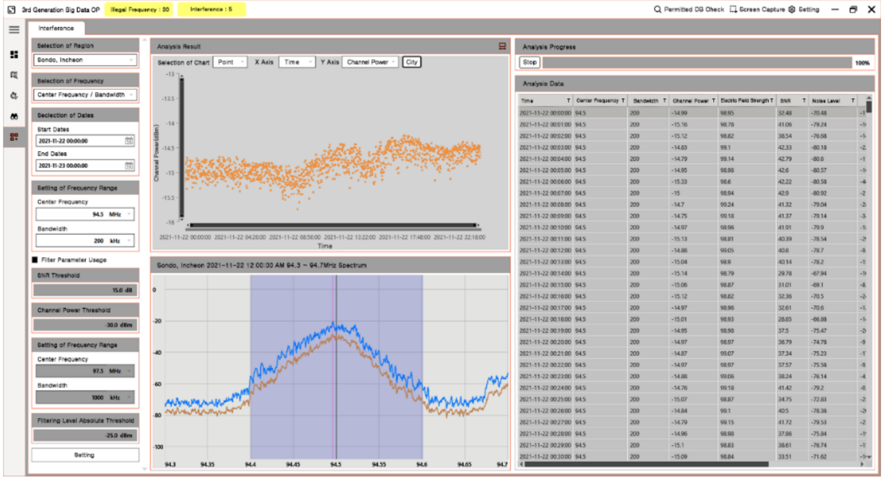


FIGURE A1-6

Raw spectrum data analysis by drill-through method based on big data



The big data spectrum monitoring solution can visualize the spatial distribution of the spectrum occupancy, radio noise, illegal frequency, and intensity of the spectrum. Figure A1-5 shows the spatial distribution of these four parameters, each plotted on a geographic map.

The big data spectrum monitoring system can analyse various parameters based on raw spectrum data, including electric field strength, channel power, noise level, etc. Administrations can access the raw spectrum data by a drill-through method which analyses the raw spectrum data (in the spectrum DB) to gain additional insights. Figure A1-6 illustrates this function.

The data-driven AI spectrum monitoring solution uses state-of-the-art machine learning technology such as an anomaly detection AI algorithm. The AI spectrum monitoring solution can observe and detect unusual spectral images and events by using an AI training method which examines spectrum images based on training results. The AI training methods include supervised, unsupervised and reinforcement learning. The AI spectrum monitoring solution typically uses the supervised learning method for analysing the spectrogram images.

In summary, the data-driven big data spectrum monitoring solution can be realized through a dedicated platform that measures the important spectrum KPIs of spectrum occupancy, spectrum quality, spectrum noise, and illegal frequencies. Administrations can investigate important spectrum incidents based on big data analysis.

There are clear advantages of the data-driven big data spectrum monitoring solution:

First, a big data spectrum monitoring solution can deal with large amounts of spectrum data. The spectrum data is collected and analysed 24/7. It is autonomously connected with the national license database for processing without human intervention. It is very hard to monitor these large amounts of data by a conventional approach.

Second, it is possible to anticipate future spectrum events by using data-driven big data spectrum monitoring to compare past spectrum data to the present. Administrations can predict future spectrum events based on spectrum big data analysis. Using this big data analysis, administrations can supervise predictable spectrum interference based on previous historical data trends.

Annex 2

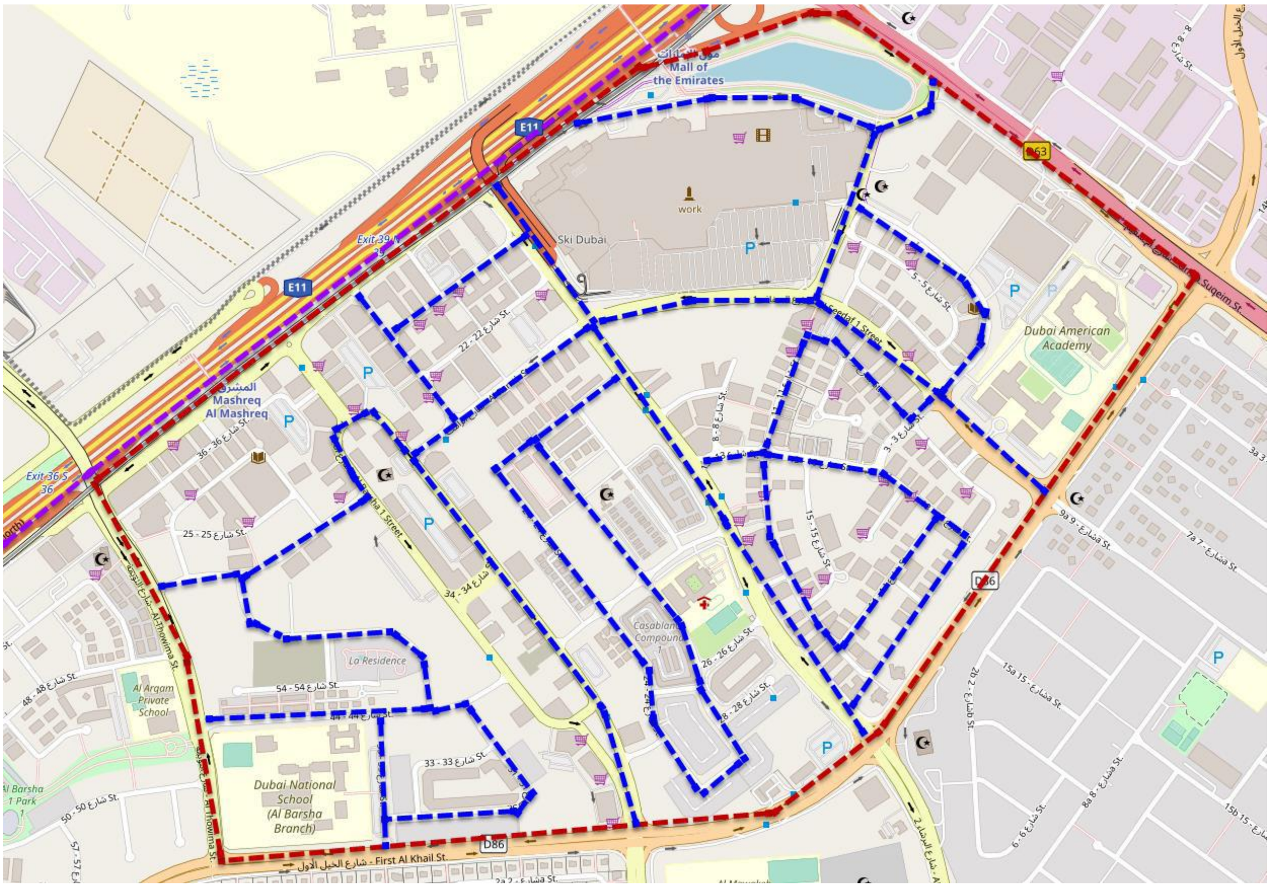
Mobile public transport based big data acquisition for spectrum mapping

For spectrum managers implementing dynamic spectrum sharing, it is important to have a realistic overview of spectrum coverage (downlink) and spatial spectrum usage (uplink). This can be achieved by collecting spectral data in many different locations across an area of interest (e.g. urban areas) and plotting the data on a map display. Spectrum mapping creates a continuous heatmap of spectrum coverage and usage and requires spatial interpolation algorithms. The better the density of the collected data, the more accurate the interpolation algorithm can predict power levels between the measurement points. To achieve this data density in urban areas, the use of public transport and service vehicles like busses, trams, taxis, garbage trucks etc. to carry RF measurement equipment is ideal. These vehicles are normally operating around the important areas of the city with a route that supports collection of dense spectral data which could save administrations costs and manpower. Further, the data density in terms of location and time on public transport cannot be achieved by a few dedicated monitoring vehicles.

Figure A2-1 illustrates the data density consideration. The area of interest is surrounded by the dashed red line. The lowest accuracy is achieved if driving only along the purple route adjacent to E11. Medium accuracy is achieved by driving the red route following E11, D63 and D86, then turning right up to E11. The highest accuracy and data density is achieved by taking the blue route following the interior of the area in addition to the red route.

FIGURE A2-1

Alternate driving routes to achieve different spectrum measurement densities



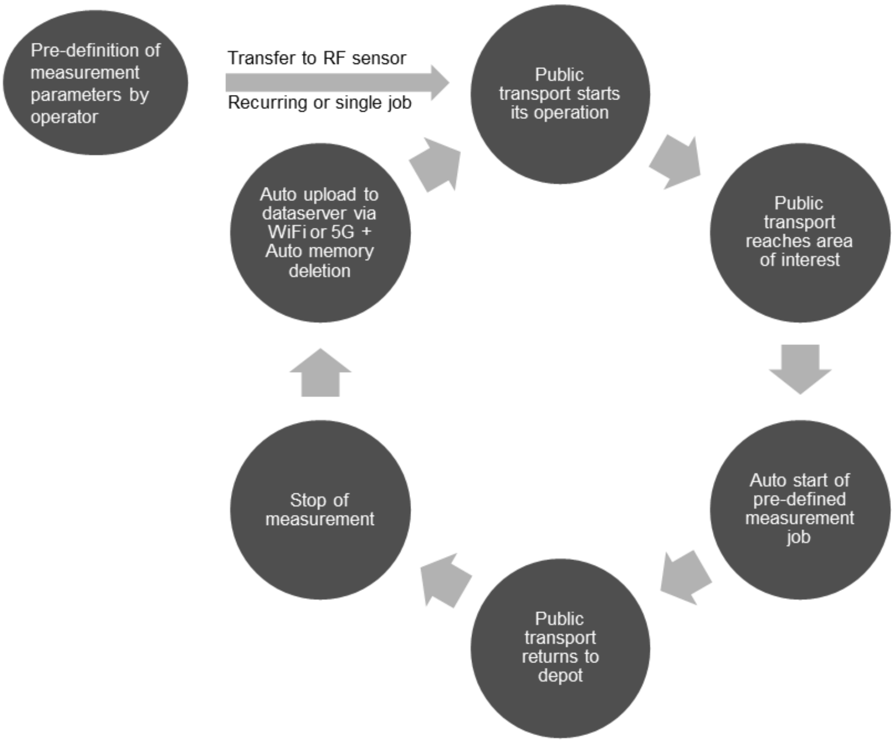
The RF monitoring sensor and antenna for public transport applications should have the following characteristics:

- Compact format;
- Embedded mass storage;
- 12-16 VDC input, to be powered by vehicle;
- Embedded backup battery in case vehicle power is lost;
- Embedded cellular networking router (LTE, 5G, Wi-Fi);
- Compact wideband antenna.

Beside the hardware requirements software automation is key for this application. The data collection and upload must be fully automated and is described in the following process circle (Fig. A2-2).

FIGURE A2-2

Spectrum measurement process for mobile monitoring platforms



The process starts with the definition of the measurement job parameters. A measurement job is a set of parameters and features to be executed by the RF sensor if the defined conditions (e.g. time and area) are met. As cell sizes and transmitting power is decreasing and the mapping requires a good spatial resolution to be meaningful, the goal is to collect as many scans per time interval while driving. It is recommended to complete one complete scan, including re-tuning time with a spatial resolution of at least 100 m. Expressed in other words, for each 100 m of driving, a band or channel of interest is re-visited. This requires high-speed monitoring receivers. The minimum tuning speed to fulfil the 100 m requirement depends on driving speed, frequency range to be monitored, resolution bandwidth (RBW), and trace averaging settings. Table A2-1 provides required re-visit times and scan speeds for an example frequency range and different driving speeds to fulfil the 100 m spatial resolution requirement. The scan speed includes time for re-tuning the receiver, as continuous scanning is required. If a frequency list scan is used, the RF sensor must be able to complete the list within the re-visit time for the next scan to start.

TABLE A2-1

Scan speed needed for 100 m resolution by drive speed

Drive speed (km/h)	Required re-visit time	Required scan speed over example range 80 MHz - 6 GHz
30	12 s	494 MHz/s
50	7.2 s	823 MHz/s
80	4.5 s	1.32 GHz/s
100	3.6 s	1.64 GHz/s
120	3 s	2 GHz/s

For a standard scan, the RBW needs to be chosen in relation to the smallest signal bandwidth expected in the frequency range of interest. For channel list scan, an appropriate RBW per channel can be selected. A list scan saves disc space but can increase the re-visit time in certain scenarios.

While the public transport drives around the city area, the received signal strength at the monitoring receiver is varying significantly. When driving near strong emitters, the receiver can get into saturation/overloading. For this reason, gain and attenuation settings are very important. Monitoring from a public-transport vehicle results in close proximity to many emitters along the route and therefore, a planned reduction in the sensitivity of the RF sensor (by adding input attenuation) provides some immunity against overloads caused by strong signals.

The amount of data per RF sensor can range from a few hundred gigabyte up to a terabyte of data per day. Field trials from 6.5 hours of data collection in a frequency range from 80 MHz up to 6 GHz produced a data amount of 300 GB. To save disc space, automation of the record start and stop times is needed. This can be achieved using an area vector-based trigger (which relies on the GPS location of the vehicle) or vehicle ignition-based trigger.

The collected data is transferred to a data server to process a map of the spectrum together with other measurements from other vehicles or fixed sensors. The data transfer can be achieved by installing a Wi-Fi or 5G access point in the depot or garage of the vehicles so that as soon as the router of the RF sensor connects to the access point, it will start its data transfer. During operation, it is recommended only to stream the trace of the live spectrum as well as the health and operational status of the RF sensor to a control centre instead of transferring the raw measurement data via the network as it will use a lot of cellular bandwidth.

As the embedded storage of the RF sensor is limited, it is suggested to automate data deletion after upload to the server is completed.

If a large number of public transports are equipped with RF sensors, a good time resolution in urban busy areas can be achieved. The more often vehicles equipped with RF equipment drive in the same areas using the same or similar routes, the better spectrum usage trends over time can be identified.

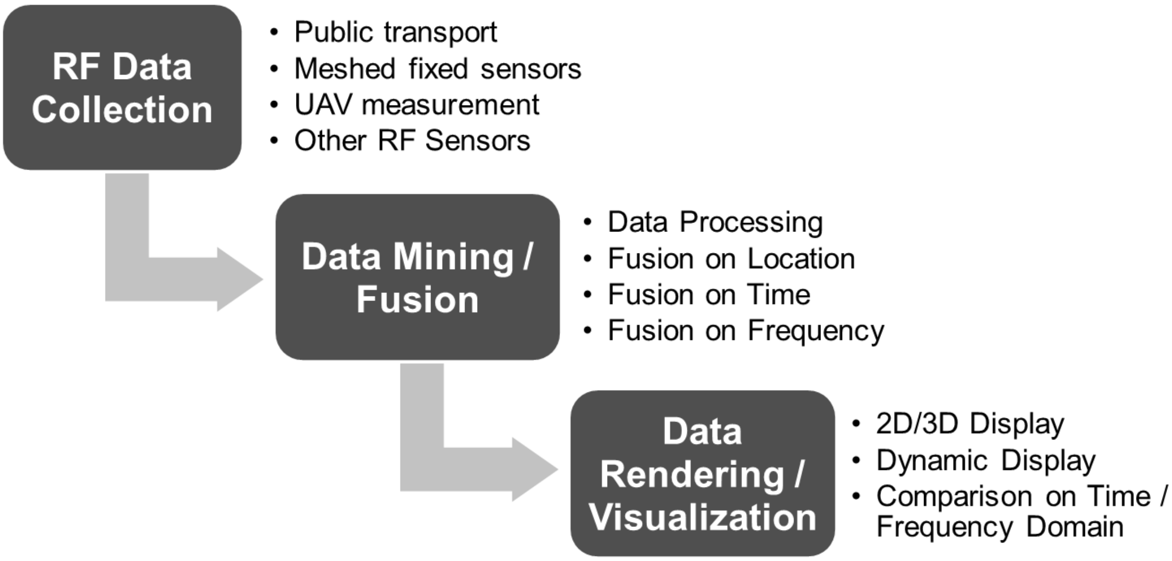
As mentioned above, big data from mobile RF sensors combined (optionally) with fixed stations can be used to process continuous heatmaps of spectrum coverage and usage for urban areas. After the data is collected and transferred to a (cloud) server, the data mining and fusion takes place. To make the measurements comparable, they must be converted to field strength values considering the signal path losses and antenna factors as well as measurement settings of the individual RF sensors. The antenna factor and associated uncertainty has to be characterized for each specific combination of vehicle type and antenna positioning. For each defined channel, a map is produced. To produce a continuous map, interpolation between the measurement points needs to be calculated. Interpolation can range from a simple linear interpolation to one supported by wave propagation models. If high precision maps with terrain and clutter data combined with a wave propagation model are used for the interpolation, the field strength inside the measurement gaps (where no vehicle was driving) can be estimated in the most accurate way. Besides geographical interpolation, time aggregation also needs to be done. Depending on the time accuracy requirement for the evaluation, this can range from a daily average, which means all collected data of a day is aggregated, up to 10 minutes time resolution in busy areas. The time resolution depends on the geographical re-visit time of public transport and fixed stations in the measurement area.

After the data mining and fusion, the data rendering/visualization needs to offer meaningful maps for spectrum managers to evaluate the real-world spectrum situation. The basic map includes a field strength distribution with colour coded dBμV/m values. The user can visualize the field strength map for all pre-defined channels and the band maps that will display the aggregated field strength for an entire band. See Fig. A2-3 for a process flow for creation of a spectral map.

If the spectrum mapping tool is connected to the transmitter database, licensed stations can be displayed together with the heatmap. By correlating the maps with licensed station information, the operator can identify hotspots which are not assigned to a station.

FIGURE A2-3

Mobile spectrum measurement workflow



The field strength maps (Fig. A2-4) can provide:

- Coverage mapping
  - Aggregated view for total band to evaluate total availability;
  - Customized views for different providers, which provide best coverage in which area? Available spectrum?
  - Comparison to wave propagation calculations, is the calculated coverage like the measured coverage?
- Frequency usage:
  - Mapping the uplink;
  - Evaluate areas of high and low usage, identify overloads;
  - Identify and localize interference or illegal spectrum usage;
  - Evaluate potential and built up a data foundation for spectrum sharing.

The big data system should offer further maps based on the field strength maps. This can be binary coverage maps, where the user can define a threshold for which the areas above are coloured as covered and below as not covered (see Fig. A2-5). A further tool for big data spectrum mapping are radiation maps in V/m. By this administrations can verify if radiation limits are met in all areas. Such maps are also well suited to be published to eliminate concerns for health problems from radiation.

FIGURE A2-4

Field strength maps

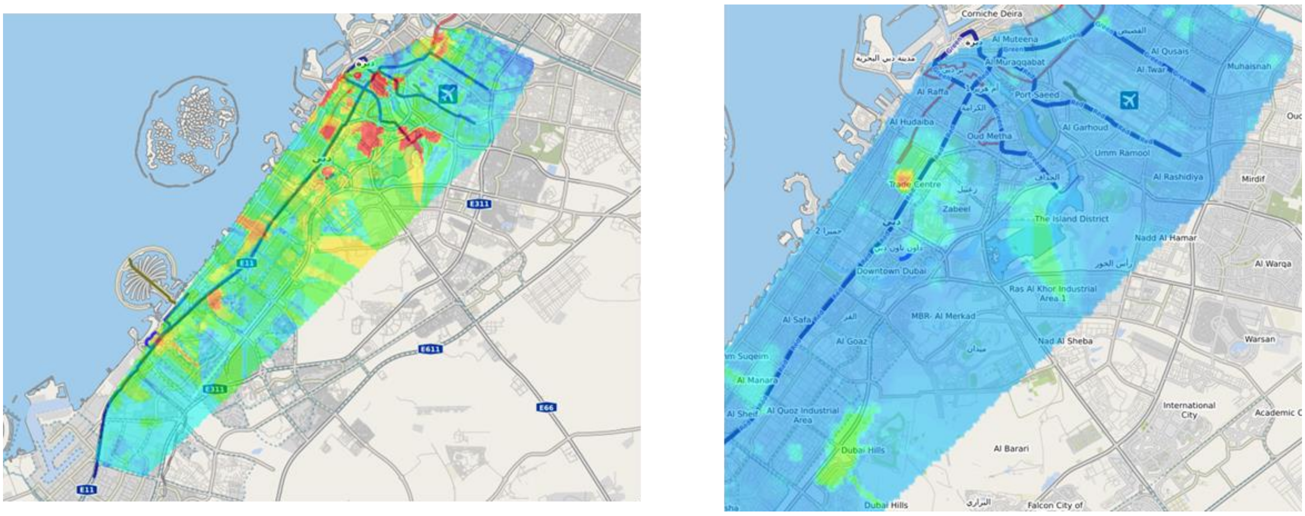


FIGURE A2-5

Coverage map