UDC 621.3:519.2

Algorithm for Spectrum Sensing and Signal Selection by External Parameters

Buhaiov M. V.

S. P. Korolov Zhytomyr Military Institute, Zhytomyr, Ukraine

E-mail: *karunen@ukr.net*

For modern radio monitoring, a panoramic view of a wide frequency band and signal selection is its most important part. The constant growth of the number of radio electronic devices and the expansion of the instantaneous bandwidth of analysis in modern radio receiving devices leads to the fact that a significant number of analog and digital signals can be observed at the same time. Automatic adaptation of radio monitoring system to further signal processing is possible due to preliminary signal selection. The goal of this research is to develop an algorithm for signals selection in panoramic radio monitoring systems by their external parameters. The essence of proposed algorithm is to detect occupied bands of radio frequency spectrum, estimate center frequency and bandwidth of each channel, noise level and signal-to-noise ratio. Creation of frequency channels allows for signal filtering and estimation of pulse durations, as well as occupancy of each channel. Estimates of parameter for each signal fragment and frequency channel are recorded in associative arrays, which simplifies further signal selection. Due to variability of noise and propagation channel, estimates of signal parameters for each signal fragment are random variables. To obtain reliable estimates of signal center frequency and bandwidth, they are further grouping. Array of data can be accessed both by frequency channel number (table rows) and by signal parameters (keys), which are table column headers. Associative relationships between data provide flexible signals filtering by any combination of parameters. To test developed algorithm, we analyzed frequency band of 933-953 MHz and used the DataFrame Multi Index data container of Pandas package of Python programming language. This structure provides multi-level indexing, flexible access to data, and a wide range of tools for their processing and modifying. Developed algorithm can be used in existing and future radio monitoring systems for radio electronic devices identification and databases creation.

Keywords: radio frequency spectrum; signal selection; external parameters; radio monitoring; associative array

DOI: 10.20535/RADAP.2024.95.5-15

Introduction

The rapid growth of electronic devices leads to a more complex electronic environment [1]. Spectrum sensing of a wide frequency band and signal selection is the most important part of modern radio monitoring. Expanding the width of the instantaneous bandwidth of modern radio receivers allows analyzing bands of radio frequency spectrum (RFS) in which a significant number of analog and digital signals can be simultaneously present. Signal channelizing allows operation in a smaller frequency band, which leads to an increase in signal-to-noise ratio (SNR) and improved processing quality. Multi-channel signal processing in panoramic radio monitoring systems requires significant computing power. Therefore, the preliminary selection of signals by external parameters allows the radio monitoring system automatically adapt to further processing, in particular, to select the algorithm for type of modulation recognizing.

1 Related works

Recently, data science technology, such as machine learning, neural networks, and big data analysis have been increasingly used to solve signal selection problems. In [2, 3], to select signals with some types of modulation (ASK/FSK), the fast Fourier transform (FFT) of the signal is calculated with the subsequent use of classifiers: decision trees, nearest neighbor methods and support vector machines with different kernels, artificial neural networks. In [4], for signals search in a wide frequency band in the presence of other signals, a high-resolution spectrogram was calculated. The information from this spectrogram is fed to a convolutional neural network with deep learning through a sliding window. In [5], a convolutional neural network is used to classify signals, and features for recognition are extracted from the spectral correlation function. In [6], the support vector method is used for signal classification. A multistage method of analyzing RFS using convolutional neural networks to SNR estimate, depending on which the further processing method is selected, is proposed in [7–9]. An approach using machine learning and entropy values as a feature vector is presented in [10]. In [11], the Akaike information criterion is used to analyze the RFS, and in [12], big data algorithms are used to build a large-scale cognitive satellite radio system. In [13], architecture for signal processing using big data and deep learning technologies is proposed. The approach is based on representation of external and internal signal parameters.

However, if the data for machine learning are of poor quality (signal fading, distortions in receiver), machine learning methods can select these uninformative parasitic characteristics as features and use them for further work.

Implementation of deep learning algorithms requires a sufficiently large amount of data for training (signal samples under different conditions), as well as sufficient computing complexity. For autonomous sensors, the latter factor can be critical. Also, at receiver location, the signal can be distorted due to channel influence (frequency-selective fading). Moreover, the influence of these factors for different channels is different and difficult to predict, which will complicate creation of a data set for training. Incomplete data set during training leads to an increase in the error rate when working with real signals. The reviewed papers describe the use of information technologies for certain stages of RFS analysis or for detecting specific types of signals.

2 Problem statement

The aim of the study is to develop an algorithm for signals selection in radio monitoring systems by their external parameters.

3 Algorithm for signal selection

The main task of preliminary (fast) analysis of RFS is to detect and select signals of radio electronic devices (RED). To do this, the following tasks must be solved:

1) frequency channels detection;

2) SNR estimation in detected frequency channels;

3) analysis of time-frequency-structure of the signals;

4) signal selection by their external parameters.

These stages make it possible to select signals of interest from the entire signal stream and implement its' detailed analysis to determine their internal parameters with the subsequent identification of the RED.

In this paper, signal selection is understood as the separation of input signal stream into frequency channels and signals extraction with specified characteristics. Signal selection is based on physical differences in frequency and temporal characteristics of the signals.

It is recommended to start the identification of digital signals with a general selection process based on external parameters [14]. External signal parameters include: center frequency, signal bandwidth, duration (for pulse signals), and spectrum shape. Another useful parameter for signal processing is SNR value in each frequency channel, which allows determining the errors of parameter estimates and the probability of correct classification using methods with known characteristics.

The essence of the proposed approach to signal selection is to detect occupied channels, determine center frequency and bandwidth of each channel, and estimate noise level and SNR. Channelizing allows for signal filtering and estimation of pulse durations and occupancy of each channel. Signal is processed in time windows, and due to the influence of random factors, estimates of parameter values for each time fragment differ from each other. Therefore, the measured parameters values for each signal fragment are recorded in associative arrays, which further facilitate signal selection. Data can be accessed both by frequency channel number (table rows) and by signal parameters (keys), which are the table column headers. Associative relationships between data provide flexible filtering of signals by any combination of parameters.

Model of received signal mixture by one antenna can be represented as follows:

$$x(t) = \sum_{i=1}^{K} s_i(t) * h_i(t) + \xi(t)\,, \tag{1}$$

where K – number of radio emission sources;

$s_i(t)$ – signal of i-th radio source;

$h_i(t)$ – impulse response of propagation of i-th channel;

$\xi(t)$ – white Gaussian noise.

Block diagram of spectrum sensing and signal selection algorithm is depicted at Fig. 1.

In block 1, the vector of parameters for spectrum sensing algorithm is entered:

1) $x(t)$ – complex samples of the received signal (also values of the sampling rate F_s and central frequency of the receiver tuning);

2) Welch periodogram parameters: N_{FFT} – FFT window length, R – overlap between windows, M – number of accumulated periodograms, w – window function;

3) L – length of moving average window for smoothing power spectral density (PSD) samples;

4) P_{F1}, P_{F2} – values of false alarm probabilities for threshold processing of test statistics (coefficient of variation) and for PSD samples, respectively;

5) $\mathbf{m_Q}, \sigma_\mathbf{Q}$ – vectors of mean and standard deviation (SD) values of coefficient of variation for the PSD noise samples for different values of Welch periodogram parameters;

6) F_{min}, F_{max} – minimum and maximum values of signal bandwidth for signal selection;

7) P_{min}, P_{max} – minimum and maximum level of signal power over noise level in frequency domain for selection;

8) f_b – channel center frequency shift.

The vast majority of parameters can be set in advance and do not need to be adjusted during the algorithm implementation. These parameters are purely technical and, if selected correctly, they do not significantly affect the algorithm. Time, frequency, and amplitude parameters can be used to search for signals of specific RED.

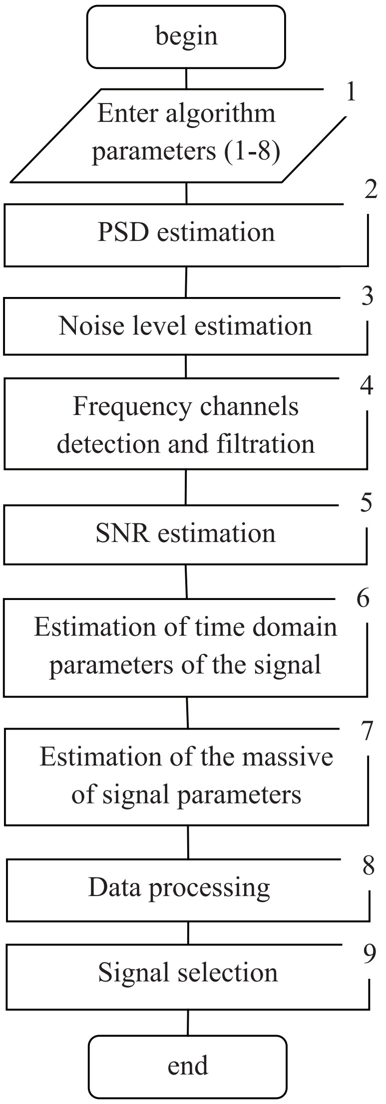


Fig. 1. Block diagram of signal selection algorithm

In block 2 PSD is calculated using the Welch periodogram. It is recommended to choose the length FFT window N_{FFT} equal to 4096 or more, since in this case at low SNR, the probability of detecting frequency channels increases. Also such signal fragment has a sufficient length in time domain to estimate signals' temporal parameters.

In block 3, the noise level in frequency domain is estimated.

In block 4, frequency channels are detected and signals are pre-filtered by bandwidth and power level to reject noise emissions.

For different types of signals, specialized methods are used to estimate carrier frequency and bandwidth. At the stage of signals selection by external parameters, it is sufficient to have their relatively rough estimates. When using digital spectrum analyzers to estimate signals frequency parameters, it is advisable to use samples of energy spectrum.

As a result of processing RFS, values of frequency channel boundaries and values of spectral samples within these boundaries are obtained. Moreover, frequency boundaries of channels f_H and f_L (k_{min} and k_{max} samples numbers) are determined as a result of threshold processing at a level close to noise level. For the sampling rate F_s, carrier frequency estimate can be obtained by the following expression:

$$f_0 = (f_L + f_H)/2 = F_s\,(k_{\max} + k_{\min})/(2N_{FFT})\,, \quad (2)$$

and signal bandwidth is calculated by the following formula

$$\Delta f = f_H - f_L = F_s\,(k_{\max} - k_{\min})/N_{FFT}\,. \quad (3)$$

This is the simplest and most intuitive way to obtain estimates of frequency parameters. In [15], it is recommended to use the β% method and the "x-dB" method to measure the width of the occupied frequency band. To apply the first method, the frequency resolution should be at least 0.03 of signal bandwidth, and SNR should be at least 30 dB. For the second method, a reference level is selected, usually 0 dB. The value of x depends on radiation class. It is recommended to measure the bandwidth at -26 dB and apply a conversion factor in accordance with the radiation class.

The β% method allows measuring the bandwidth of signals regardless of the modulation type. However, it is advisable to use it to estimate bandwidth of digital signals at low SNR values. In cases of interference, the "x-dB" method is more appropriate. In this case, to obtain reliable estimates of bandwidth, signal spectrum should contain 100-200 spectral samples [16].

In [17], to estimate signals frequency parameters the samples of PSD are used. Taking into account values of signal PSD samples P_{xx}, its central frequency is proposed to calculate by the following expression:

$$f_0 = \frac{F_s}{EN_{FFT}} \sum_{i=k_{\min}}^{k_{\max}} iP_{xx}\,[i], \quad (4)$$

where E – signal energy; i – number of signal bin.

Signal bandwidth (effective) can be calculated as follows:

$$\Delta f = \frac{1}{E}\sqrt{\frac{F_s}{N_{FFT}}\sum_{i=k_{\min}}^{k_{\max}} i^2 P_{xx}^2\,[i]}. \qquad (5)$$

The estimates obtained in accordance with (5) differ from signal bandwidth at the noise level (3) and depend on the shape of signal spectrum. This creates prerequisites for signal selection using test statistics from spectral samples.

In most digital radio systems signal bandwidth remains constant over time because data stream is transmitted at a constant symbol rate. In analog systems, especially when transmitting audio signals (F3E, A3E, J3E), instantaneous bandwidth changes rapidly. Also, due to the variability of noise at low SNR, each realization of signal spectrum is represented by a different number of frequency samples, which leads to fluctuations in values of frequency parameters. Therefore, in order to obtain reliable estimates of center frequency and signal bandwidth, it is necessary to perform additional processing of sequences of obtained parameter values.

In block 5, a vector of SNR values is generated for each channel.

In block 6, frequency channels are filtered in time domain, pulses are detected, and their temporal parameters are estimated. Detected in each channel pulses with probability of false alarm P_{F3} are filtered by duration. Since several pulses can be detected in frequency channel during signal analysis interval, a vector of $\mathbf{T}$ values is formed. Estimate of channel occupancy is calculated as the ratio of samples number that exceeded threshold and were filtered by duration as signal samples to the total number of signal samples. Experimental studies have shown that shift of center frequency in channel f_b has virtually no effect on the quality of signal detection in time domain.

The number of samples for which signal is filtered in channels is $N = (M-1)\cdot R + N_{FFT}$. Filtering is performed by calculating the inverse FFT of complex frequency samples of length N. Prior to this, all frequency samples outside the specified channel are equalized to zero.

In block 7, an associative array is formed from measured values of signal parameters. The structure of array for storing information on results of RFS analysis is given in Table 1.

Main parameter in Table 1 is number of frequency channel to which the rest values are assigned. This approach facilitates further filtering of channels by values of other parameters. In Table 1, number of frequency channels is indicated by the line number. Number of detected frequency channels for each periodogram can be different. Therefore, for different periodograms, the same frequency channel may have a different number, since channel numbering each time starts with the lowest value of center frequency nominal. This is due to fact that signal can be pulsed or with a frequency hopping and for different signal fragments this channel can be busy or free.

Block 8 processes tabular data, reassigns channel numbers, and generates descriptive statistics: average f_0 and SD σ_{f0} values of center frequency, channel bandwidth $\Delta f, \sigma_{\Delta f}$, SNR, σ_{SNR}, occupancy η, σ_η, and pulse duration τ.

Due to noise influence, variability of signals, measurement errors for each signal fragment, measured parameter values are random values. Therefore, when processing an array of measurements, it is necessary to obtain estimates of channel center frequencies. To do this, we accumulate vector of measured values of center frequencies $\mathbf{f_0}$ for all channels into groups by defining the boundaries of these groups and calculating the center frequencies of channels as arithmetic averages within the groups. Vector of i-th group boundaries is calculated by the following expression:

$$\mathbf{f_i} = \arg\left[diff\left(sort\left(\mathbf{f_0}\right)\right) \geq \Delta F\right], \qquad (6)$$

where ΔF – minimal separation between frequency channels.

Table 1 — Array structure with results of RFS analysis

Periodogram number	Frequency channel number	f_0, MHz	f_L, MHz	f_H, MHz	SNR, dB	$\mathbf{T}$, ms	η,%
0	0	933.289	933.161	933.417	6.44	[0.56, 0.19, …]	35
	1	933.770	933.565	933.975	5.31	[0.57, 0.54, …]	41
	…	…	…	…	…	…	…
1	0	933.295	933.176	933.415	4.86	[0.55, 0.57, …]	22
	1	933.770	933.558	933.983	6.88	[0.56, 0.56, …]	51
	…	…	…	…	…	…	…
2	…	…	…	…	…	…	…

Estimation of lower f_L and upper f_H values of frequency channel boundaries is calculated using the vectors of measured $\mathbf{f_L}$ and $\mathbf{f_H}$ values for all signal fragments. The lowest frequency limit value for i-th channel is calculated according to the following expression:

$$f_{Li} = mean\left(\mathbf{f_L}\left[iP_i\colon(i+1)\,P_i\right]\right), \quad i=0,1,\ldots,N_{ch}-1, \tag{7}$$

where P_i is the number of signal fragments of i-th channel in which signal was detected; N_{ch} – number of frequency channels.

Value of the upper frequency limit of a channel is calculated in the same way. Bandwidth of each channel is calculated as the difference between its high and low frequencies.

After determination of signal duration and detecting all frequency channels, each channel is assigned a sequence number associated with center frequency value in ascending order.

After all the data are processed, arrays of channel center frequencies and the width of each channel are formed, the channel numbers are reassigned. The reassignment is necessary because in each signal fragment the number of channels can be different or the same with different carrier frequency nominal. Number of frequency channels for the entire signal can be greater than for each individual signal fragment. Reassignment provides unification of channel number with its center frequency and other parameters.

In each new channel, for all signal fragments, we get an array of values filtered by pulse duration. One channel can have several signal starts and ends within the analyzed time fragment. Moreover, signal can start in the previous fragment and end in the next one.

Due to noise and fading in propagation channel, number of detected pulses with different durations can be quite large. Assuming that the RED uses a limited set of pulse durations, the task of finding their values consists in forming an empirical probability density of these durations and finding those values that correspond to its maximums. For a given vector of pulse durations $\mathbf{T}$ in frequency channel, histogram calculation results in vectors of centers of partitioning intervals $\mathbf{b}$ and the number of values falling into each interval $\mathbf{a}$.

Then duration of k-th pulse can be found by the following expression:

$$\tau_k = b_k\left[\arg\max\left(a_k\right)\right], \tag{8}$$

where k also corresponds to the number of the histogram maximum.

To find all the maxima of a histogram, it is needed to set search criteria by the value of maximum and distance between them. It is possible create special criteria for a particular case.

Average utilization of the i-th channel in time domain can be calculated using the following expression:

$$\eta_i = \frac{P_i}{D}\cdot\frac{F_s\sum\limits_{k=1}^{P_i}\tau_{ik}}{(M-1)\,R+N_{FFT}}, \tag{9}$$

where D – number of analyzed signal fragments; P_i – number of pulses in i-th channel; τ_{ik} – duration of k-th pulse in i-th channel.

Block 9 selects (searches) signals in accordance with the specified criteria. The result of previous blocks of algorithm can be presented as an associative table (matrix $\mathbf{S}$), each row of which corresponds to the number of reassigned frequency channel, and the columns contain values of center frequency, bandwidth, pulse duration, occupancy and SNR. Logical processing of values from associative table (matrix $\mathbf{S}$), taking into account frequency and temporal parameters, allows signal selection and association of individual frequency channels with RED, which uses frequency hopping. Signal selection (search) is performed using the condition vector $\mathbf{C}$, which contains restrictions on values of signal parameters. When searching for a particular RED, it is recommended to form the vector $\mathbf{C}$ from double inequalities, since due to influence of random factors, measured values of parameters may differ from true ones. The result of selection $\mathbf{S}[\mathbf{C}]$ is a new matrix with fewer rows, cell values of which fulfill filtering requirements.

Developed algorithm for spectrum sensing and signal selection requires a large number of input parameters for its operation. That's why, it is possible to flexibly adjust this algorithm to work with different types of signals without changing its structure, but only by adjusting some parameters.

Values of parameters depend on analyzed frequency range and a priori information about electronic environment. For example, in frequency range below 100 MHz, signals with a bandwidth of more than a few hundred kHz are practically not found.

4 Results and discussion

Analyzed signal was recorded using software-defined transceiver HackRF One in 20 MHz bandwidth (933-953 MHz). The values of algorithm parameters used in the study are given in Table 2.

Table 2 — Algorithm parameters

Parameter	Value
N_{FFT}	16384
R	0,5 N_{FFT}
M	30
w	Hamming
L	40
P_{F1}	0,1
P_{F2}	0,1
F_{min}	200 kHz
F_{max}	10 MHz
P_{min}	5 dB
P_{max}	60 dB
f_b	100 kHz
P_{F3}	10^{-6}
T_{min}	0,1 ms
T_{max}	1 ms

For testing the developed approach, we used the DataFrame Multi Index data container of Pandas package of Python programming language. This structure provides multi-level indexing, flexible access to data, and a wide range of tools for processing and modifying them, including adding records (columns) with new keys generated as a result of processing existing data [18].

Figure 2 shows a spectrogram of a signal recording with duration of about 3.5 s. For parameters given in Table 2, this signal recording contains D = 296 signal fragments with a length of 253952 samples each, which is about 12.7 ms. The figure shows that some signals are pulsed, while others appear to be continuous.

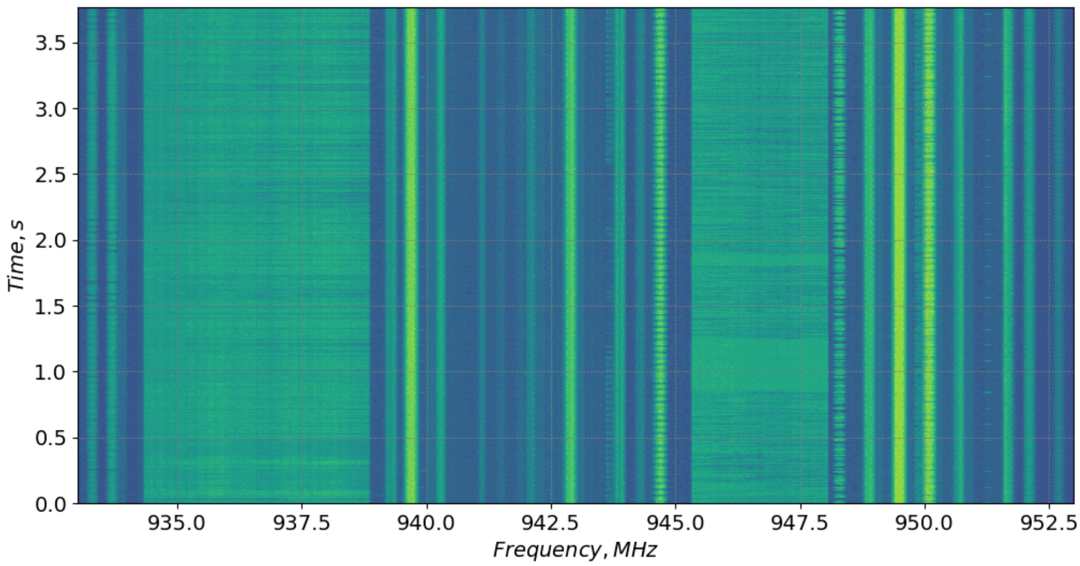


Fig. 2. Spectrogram of a signal in an analyzed frequency band

Figure 3 shows smoothed PSD in the analyzed frequency band, level noise, and threshold, as well as the selected frequency channels. After applying the power level and spectral width filters, band occupancy is about 83%.

Figure 4 shows PSD for each frequency channel after additional smoothing with a moving average window. For each spectrum its' channel number, bandwidth at threshold level and SNR are also shown. As we can see, in some channels $(1, 6, 7, 8, 12, 17)$, additional processing is required to filter out out-of-band spectral components.

Figure 5 shows time realizations of signals in selected frequency channels. From these graphs, it is possible to preliminarily assess the signals temporal structure and feasibility of further processing in each channel, taking into account SNR estimates.

As a result of signal analysis 22 frequency channels were detected in a given frequency band. As can be seen from the above figures, only 20 channels were detected for the first FFT realization (Fig. 3).

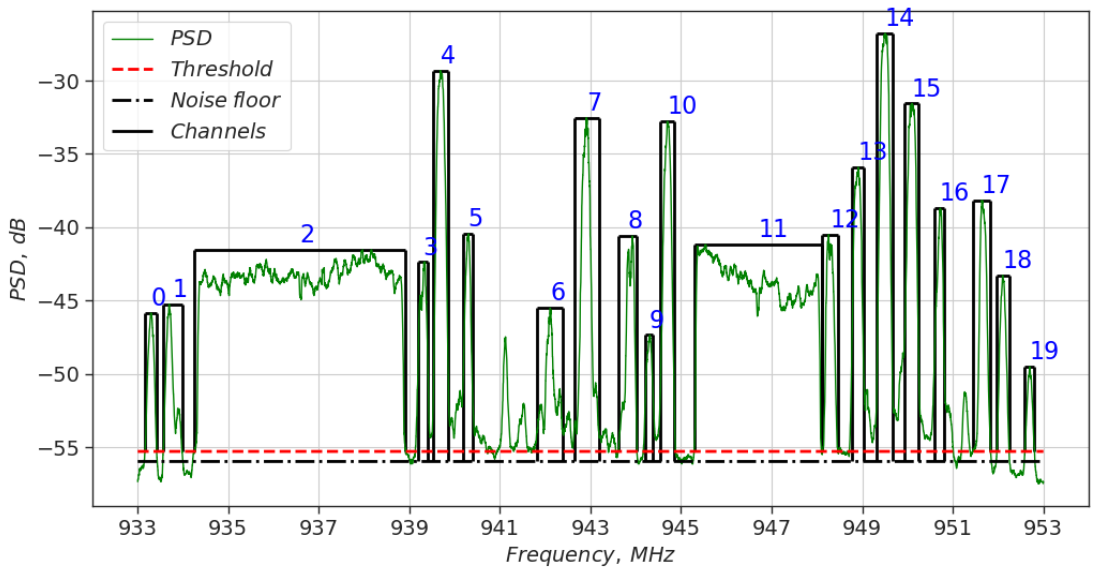


Fig. 3. Channelized spectrum

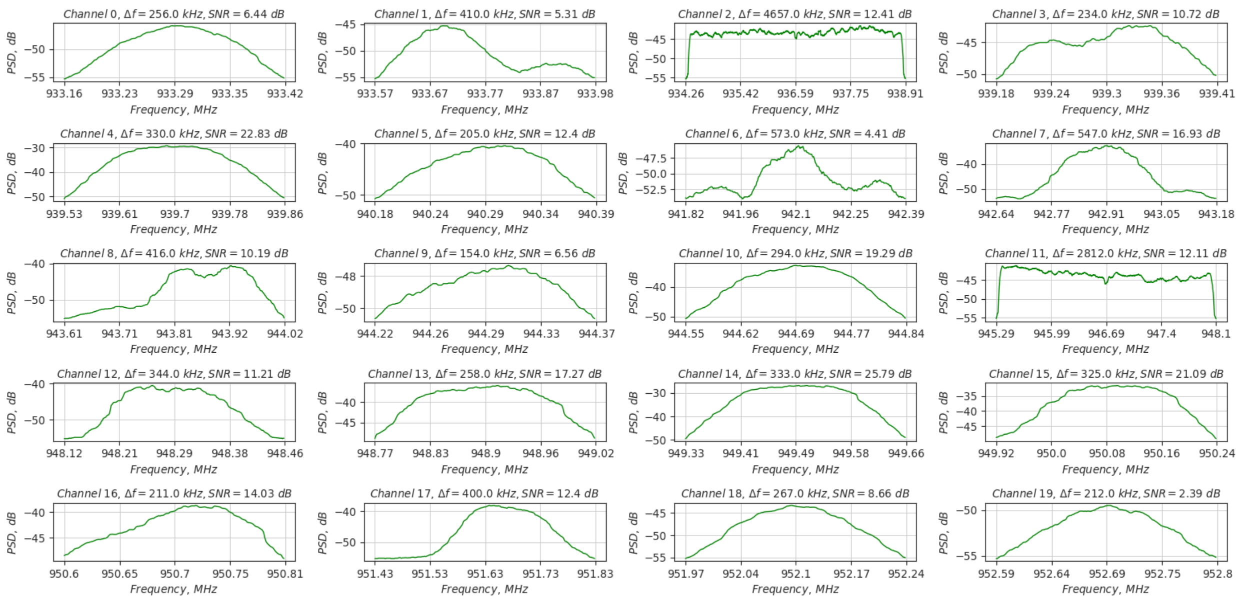


Fig. 4. PSD for each frequency channel after additional smoothing

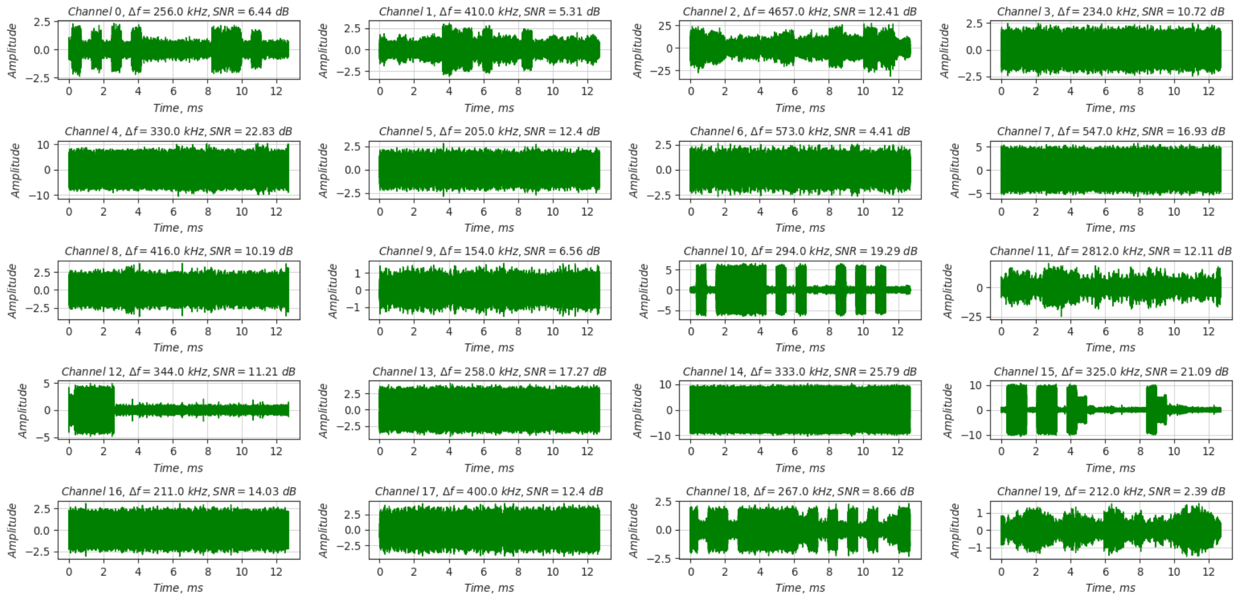


Fig. 5. Channelized signals in time domain

Figure 6 shows signal processing in time domain for channel 0. Since in process of data generation, short bursts in the time domain can be created that are not related to useful pulses, filtering by duration allows us to discard some of uninformative data about the transmission structure.

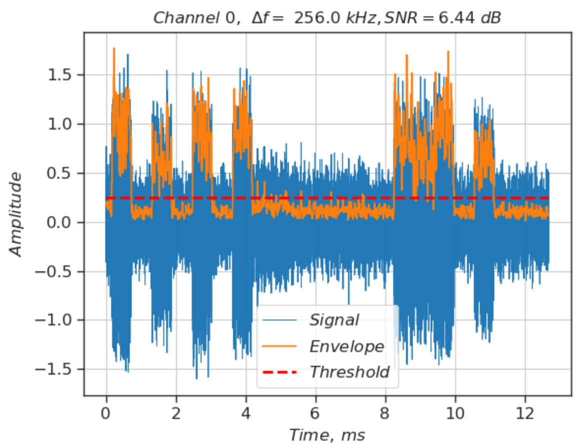
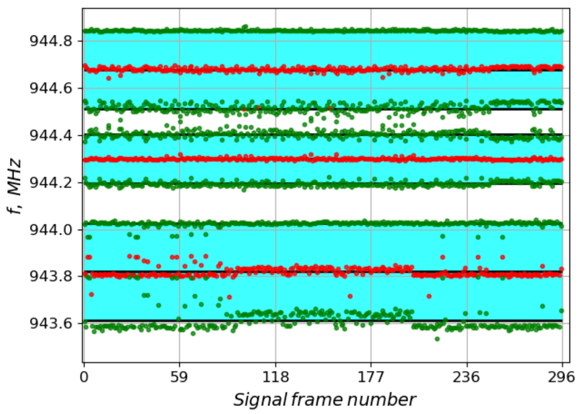
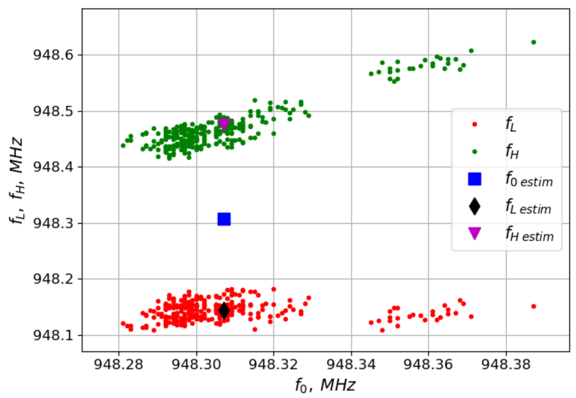


Fig. 6. Time domain signal processing for channel No. 0

Figure 7a shows measured values of signal center frequency (red dots), upper and lower frequency limits (green dots), and frequency channels as a result of processing (shaded). Figure 7b shows scatter plot of measured values of frequency channel boundaries for each measured value of center frequency, as well as averaged estimates of frequency parameters according to expressions (2-3).

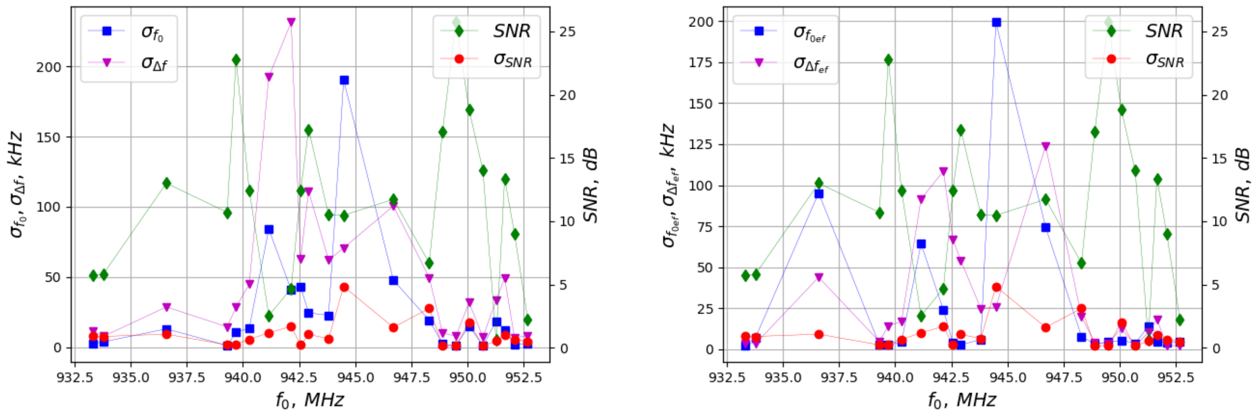


(a)



(b)

Fig. 7. Measured values of frequency parameters and frequency channels (a) and scatter plot of channel frequency limits (b)



(a) (b)

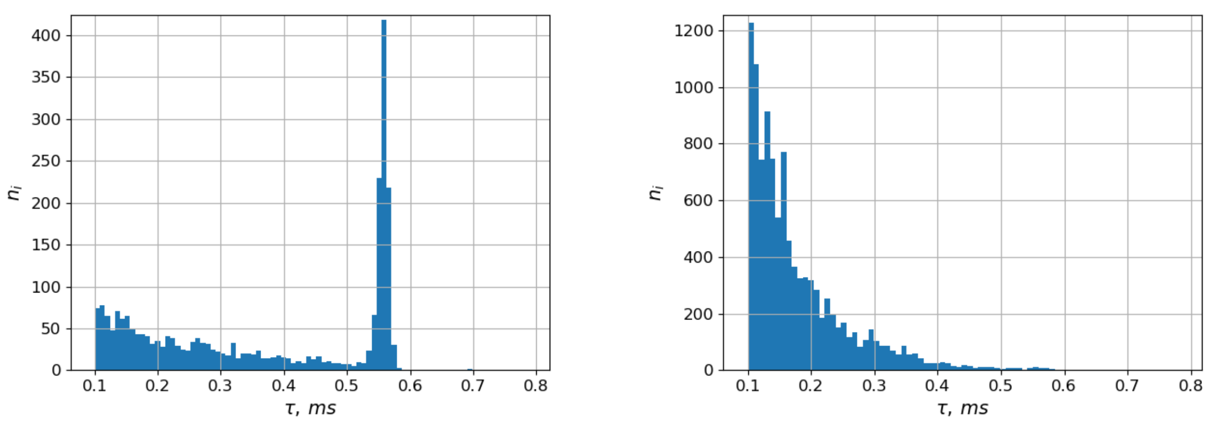
Fig. 8. Dependences of σ_{f0}, $\sigma_{\Delta f}$, SNR, σ_{SNR} estimates on center frequency (a) and a similar graph for f_0, Δf SD estimates, obtained by expressions (4) and (5) (b)

Figure 8a shows dependence on frequency of center frequency estimates SD, bandwidth, calculated by expressions (2) and (3), as well as mean values and SD of SNR. Figure 8b shows similar dependencies for which values of center frequency and bandwidth are calculated by expressions (4) and (5).

As a result of these dependencies analysis, it was found that estimates of frequency parameters obtained by expressions (4)-(5) have, on average (for 22 signals), 2 times less variance than those obtained by expressions (2)-(3). No dependence was found between the average values of SNR in channel and SD of frequency parameter estimates. However, it is worth noting that for channels with a larger SD of SNR, larger SD of frequency parameter estimates are observed. SNR value in a channel at short time intervals changes mainly due to changes in signal power, which may indicate amplitude modulation or presence of deep fading in propagation channel. In this case, SNR variability measure is its SD (Fig. 8). As SNR increase, SD should decrease. If this does not happen, it can be assumed that there are internal regularities (modulation parameters) in transmitted data.

Also, for signals with a rectangular spectrum, SD of center frequency estimates (4) and bandwidth (5) is larger than for signals with another spectrum shape and estimates obtained according to expression (2). In addition, for signals with an almost rectangular spectrum (OFDM), the ratio of estimates obtained by expressions (3) and (5) is about 1, while for signals with a bell-shaped spectrum, this ratio is about 3.

Figure 9 shows histograms of measured values of pulse durations in channel No. 0 (a) and No. 19 (b). Both histograms begin with a value of 0.1 ms, which corresponds to the set value of T_{min}. For the first case, average SNR is 5.7 dB and SNR SD is 0.9 dB, and for the second case, 13.4 dB and 1 dB, respectively. In Fig. 9a, a clearly distinguished maximum of histogram is observed, which corresponds to pulse duration of about 0.56 ms. In Fig. 9b, despite the high SNR, no pulses with a fixed duration were detected.



(a) (b)

Fig. 9. Histograms of measured pulse durations in channel No. 0 (a) and No. 19 (b)

Both histograms also show an exponential probability density of pulse durations. In this case, this can be explained by the fact that in presence of deep fading in propagation channel, signal envelope has a parasitic amplitude modulation. Moreover, amplitude of signal envelope changes randomly. In this case, time interval between two consecutive threshold crossings by signal envelope is also random exponentially distributed variable.

Figure 10 shows dependence of occupancy η, SD occupancy σ_η, and detected pulses duration on channel center frequency. In those channels where $\eta = 100\%$, $\sigma_\eta = 0$ and $\tau = 0$, a continuous signal is transmitted, level of which never falls below threshold. This is mainly observed for channels in which an OFDM signal is transmitted. As it can be seen, for most frequency channels, pulse duration is (0.54-0.57) ms, which corresponds to frame duration of the 2G mobile communication (0.577 ms).

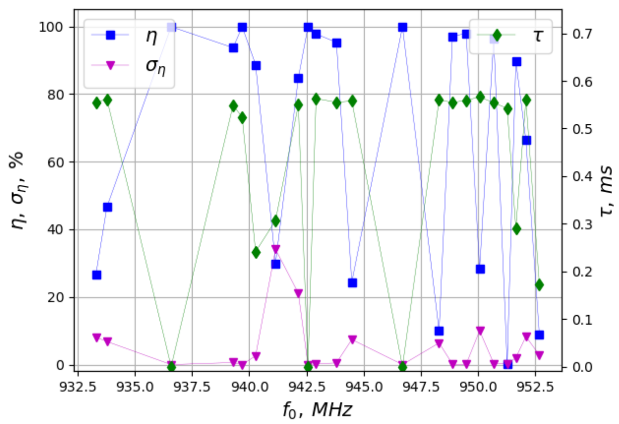


Fig. 10. Dependence of η, σ_η, τ estimates on center frequency

channel_new	f_0, MHz	BW, kHz	T, ms	occup, %	SNR, dB
0	933.292	247.4	0.55	26.0	5.7
1	933.776	426.3	0.56	47.0	5.8
2	936.586	4657.2	0.00	100.0	13.0
3	939.296	239.4	0.55	94.0	10.7
4	939.693	335.2	0.52	100.0	22.8
5	940.280	219.0	0.24	89.0	12.4
6	941.153	376.4	0.31	30.0	2.5
7	942.150	615.0	0.55	85.0	4.7
8	942.571	1671.2	0.00	100.0	12.4
9	942.926	526.8	0.56	98.0	17.2
10	943.819	410.4	0.56	95.0	10.5
11	944.489	271.9	0.56	24.0	10.5
12	946.707	2832.0	0.00	100.0	11.7
13	948.307	326.2	0.56	10.0	6.7
14	948.896	263.0	0.55	97.0	17.1
15	949.493	341.0	0.56	98.0	25.7
16	950.080	319.1	0.57	28.0	18.8
17	950.699	218.9	0.55	96.0	14.0
18	951.287	208.4	0.54	0.0	0.6
19	951.667	313.6	0.29	90.0	13.4
20	952.106	267.5	0.56	66.0	9.0
21	952.695	213.3	0.17	9.0	2.2

(a)

channel_new	f_0, MHz	BW, kHz	T, ms	occup, %	SNR, dB
1	933.776	426.3	0.56	47.0	5.8
4	939.693	335.2	0.52	100.0	22.8
6	941.153	376.4	0.31	30.0	2.5
10	943.819	410.4	0.56	95.0	10.5
11	944.489	271.9	0.56	24.0	10.5
13	948.307	326.2	0.56	10.0	6.7
14	948.896	263.0	0.55	97.0	17.1
15	949.493	341.0	0.56	98.0	25.7
16	950.080	319.1	0.57	28.0	18.8
19	951.667	313.6	0.29	90.0	13.4
20	952.106	267.5	0.56	66.0	9.0

(b)

Fig. 11. Table with parameter estimated for an analyzed signal before (a) and after filtering by channel bandwidth (b)

Figure 11a shows a Pandas DataFrame table with parameters estimated for analyzed signal, and Fig. 11b shows this table after filtering by channel bandwidth in accordance with the following condition: $250 < \Delta f < 500$ kHz. As can be seen from this table, associative relationships between channel number and rest of parameters are preserved. Similarly, it is possible to implement filtering of table rows by any parameters contained in column headers.

In the table (Fig. 11a), we can also add columns with values of other parameters and modify existing ones. For example, a column can be added in which, for given signal parameters (each row), an identifier can be recorded of a specific RED (radio transmission standard). In future, this information can be used for RED identification.

Conclusions

Modern information technologies for signal selection give new opportunities for a comprehensive and exhaustive analysis of data obtained in result of spectrum sensing. Scientific novelty of proposed algorithm lies in creation and processing of associative arrays of external signal parameters estimates. This approach provides flexible signal selection in panoramic radio monitoring systems. Developed algorithm can be used in existing and prospective radio monitoring systems for RED identification and databases creation.

Prospects for further research in this area are related to development of methods and algorithms for signals selection by spectrum shape, as well as processing filtered signals in time domain and improving methods for pulse durations estimation.

References

[1] Liang Y.-C. (2020) Dynamic Spectrum Management. From Cognitive Radio to Blockchain and Artificial Intelligence. *Springer*, 180 p. doi: 10.1007/978-981-15-0776-2.

[2] Saber M. et. al. (2020) Spectrum Sensing for Smart Embedded Devices in Cognitive Networks using Machine Learning Algorithms. *Procedia Computer Science*, Vol. 176, pp. 2404–2413. doi: 10.1016/j.procs.2020.09.311.

[3] Zhang Y. et al. (2017) A Spectrum Sensing Method Based on Signal Feature and Clustering Algorithm in Cognitive Wireless Multimedia Sensor Networks. *Advances in Multimedia*, Vol. 2017, 11 p. doi: 10.1155/2017/2895680.

[4] Franco H., Cobo-Kroenke C., Welch S., Graciarena M. (2020) Wideband Spectral Monitoring Using Deep Learning. *Proceedings of the 2nd ACM Workshop on Wireless Security and Machine Learning (WiseML2020)*, pp.19–24. doi: 10.1145/3395352.3402620.

[5] Tekbiyik K. et al. (2021) Spectrum Sensing and Signal Identification with Deep Learning based on Spectral Correlation Function. *IEEE Transactions on Vehicular Technology*, Vol. 70, Iss. 10, pp. 10514-10527. doi: 10.1109/TVT.2021.3109236.

[6] Tekbiyik K. et al. (2019) Multi-Dimensional Wireless Signal Identification Based on Support Vector Machines. *IEEE Access*, Vol. 7, pp. 138890-138903. doi: 10.1109/ACCESS.2019.2942368.

[7] Jeevangi S., Jawaligi S., Patil V. (2022) Deep Learning-based SNR Estimation for Multistage Spectrum Sensing in Cognitive Radio Networks. *Journal of Telecommunications and Information Technology*, Vol. 4, pp. 21-31. doi: 10.26636/jtit.2022.164922.

[8] Bedir O., Ekti A. R., Ozdemir M. K. (2023) Exploring Deep Learning for Adaptive Energy Detection Threshold Determination: A Multistage Approach. *Electronics*, Vol. 12, 18 p. doi: 10.3390/electronics12194183.

[9] Bari F., Agrawal P., Chatterjee B., Sen S. (2022) Statistical Analysis Based Feature Selection Enhanced RF-PUF With >99.8% Accuracy on Unmodified Commodity Transmitters for IoT Physical Security. *Frontiers in Electronics*, Vol. 3, 14 p. doi: 10.3389/felec.2022.856284.

[10] Baldini G., Chareau J.-M., Bonavitacola F. (2021) Spectrum Sensing Implemented with Improved Fluctuation-Based Dispersion Entropy and Machine Learning. *Entropy*, Vol. 23, 24 p. doi: 10.3390/e23121611.

[11] Zayen B., Hayar A., Kansanen K. (2009) Blind Spectrum Sensing for Cognitive Radio Based on Signal Space Dimension Estimation. *IEEE International Conference on Communications*, pp. 1-5. doi: 10.1109/ICC.2009.5198794.

[12] Yang M., Shao X., Xue G. et al. (2021) Big data theory based spectrum sensing algorithm for the satellite cognitive radio network. *Wireless Networks*, 9 p. doi: 10.1007/s11276-021-02808-7.

[13] Zheng S. et al. (2018) Big Data Processing Architecture for Radio Signals Empowered by Deep Learning: Concept, Experiment, Applications and Challenges. *IEEE Access*, Vol. 6, pp. 55907-55922. doi: 10.1109/ACCESS.2018.2872769.

[14] Recommendation ITU-R SM.1600-3. Technical identification of digital signals SM Series Spectrum management. (2017) *ITU*, 25 p.

[15] Recommendation *ITU-R SM.443* – Bandwidth measurement at monitoring stations.

[16] Handbook. Spectrum monitoring. (2011) *ITU Radiocommunication Bureau*, 678 p.

[17] Cook C. E., Bernfeld M. (1993) *Radar Signals: An Introduction to Theory and Applications.* Artech House, Inc.: Norwood, MA, USA. 552 p.

[18] VanderPlas J. (2017) *Python Data Science Handbook. Essential Tools for Working with Data.* O'Reilly Media. 647 p.

Алгоритм аналізу радіочастотного спектра та селекції сигналів за зовнішніми параметрами

Бугайов М. В.

Для сучасного радіомоніторингу панорамний огляд широкої смуги частот і селекція сигналів складає його найважливішу частину. Постійне зростання кількості радіоелектронних засобів та розширення ширини миттєвої смуги аналізу у сучасних радіоприймальних пристроях призводить до того, що одночасно може спостерігатися значна кількість аналогових та цифрових сигналів. Автоматична адаптація системи радіомоніторингу до подальшої обробки сигналів можлива завдяки попередній селекції сигналів. Метою дослідження є розробка алгоритму для селекції сигналів у панорамних системах радіомоніторингу за їх зовнішніми параметрами. Сутність запропонованого алгоритму полягає у виявленні зайнятих ділянок радіочастотного спектра, визначенні центральної частоти та ширини смуги кожного каналу, оцінюванні рівня шуму та відношення сигнал-шум. Утворення частотних каналів дозволяє проводити фільтрацію сигналів та оцінювати тривалості імпульсів, а також завантаженість кожного каналу. Оцінки значень параметрів для кожного фрагменту сигналу та частотного каналу записуються в асоціативні масиви, що в подальшому полегшує селекцію сигналів. Через мінливість шуму та каналу поширення оцінки значень параметрів сигналів для кожного фрагменту сигналу будуть випадковими величинами. Для отримання надійних оцінок центральної частоти та ширини смуги сигналу проведено їх додаткову обробку шляхом групування виміряних значень. Доступ до даних масиву можна здійснювати як за номером частотного каналу (рядки таблиці), так і за параметрами сигналів (ключами), які є заголовками стовпців таблиці. Асоціативні зв'язки між даними забезпечать гнучку фільтрацію сигналів за будь-якими комбінаціями параметрів. Для перевірки розробленого підходу було проаналізовано смугу частот 933-953 МГц та використано контейнер для даних DataFrame Multi Index пакету Pandas мови програмування Python. Дана структура забезпечує багаторівневу індексацію, гнучкий доступ до даних та широкий набір інструментів для їх оброблення та модифікації. Розроблений алгоритм може бути використаний в існуючих та перспективних системах радіомоніторингу для ідентифікації радіоелектронних засобів та формування баз даних.

Ключові слова: радіочастотний спектр; селекція сигналів; зовнішні параметри; радіомоніторинг; асоціативний масив