

Search for transient signals in the data of Decameter Survey of the Northern Sky

I. P. Kravtsov^{1*}, *V. V. Zakharenko*¹, *I. Y. Vasylieva*¹, *S. S. Mykhailova*²,
*O. M. Ulyanov*¹, *A. I. Shevtsova*¹, *A. O. Skoryk*¹, *P. Zarka*³, *O. O. Konovalenko*¹

¹Institute of Radio Astronomy of the NAS of Ukraine, Mystetstv St., 4, 61002, Kharkiv, Ukraine

²V. N. Karazin Kharkiv National University, Svobody Sq., 4, 61022, Kharkiv, Ukraine

³LESIA & USN, Observatoire de Paris, CNRS, PSL/SU/UPMC/UPD/SPC/UO/OSUC, Meudon, France

Present paper describes the data analysis of Decameter Pulsars and Transients Survey of the Northern Sky using UTR-2 radio telescope in order to search for pulsars and transient sources. A special attention is paid to algorithms and routines for distinguishing the signals of cosmic nature against terrestrial artificial and natural interference.

Key words: methods: data analysis – surveys – stars: neutron

INTRODUCTION

After discovery of pulsars [1] sky surveys with the purpose of searching new pulsars became one of necessary sections of radio astronomy. The result of these surveys is the discovery of various types of pulsed [3, 4] and transient radiation sources (repetitive/non-repetitive transients), which are generated by different types of neutron stars (NS), or similar sources (presumably — mergers of two NS or black hole and NS). Up to now a few dozens of such “blind” searches — surveys — have been carried out. Due to technology improvement (radio telescopes’ sensitivity and receivers’ bandwidth) such searches will keep on developing for the years to come.

Successful detection of vast majority of nearby (with dispersion measure $DM < 30 \text{ pc/cm}^3$) known pulsars [8] proved the possibility of carrying out the search for sources of pulsed and transient emission in the decametre range despite the high temperature of the Galactic background, the presence of strong radio interference (RFI) and steep increase in characteristic scattering time of pulsed signals in the interstellar medium (ISM) at low frequencies.

Since 2011 such survey is being carried out at UTR-2 radio telescope [2, 5]. Its prominent feature is high sensitivity to the dispersion delay of the signals in the ISM, which allows to distinguish between the natural signals (frequency response of the delay is proportional to f^{-2}) and artificial ones — with other dependencies on frequency, (e. g., f^{-1}).

Distinguishing the signals of cosmic nature against terrestrial RFI (both natural and artificial) is an extremely important and challenging task in

the data processing in the decameter range, which is saturated with interference. This paper is devoted to methods of searching for transient sources (e. g., rotating radio transients, X-ray dim isolated NS, anomalous X-ray pulsars, individual pulses of pulsars, etc.) in the Decameter Survey of the Northern sky data.

DATA ANALYSIS

In the previous paper [7] the properties of candidate signals were analysed. These properties were obtained by using the RFI mitigation routine, the dispersion delay eliminating (de-dispersion) routine and by frequency integration [6]. The results of processing of approximately 30-35% of the recorded data were presented in [7]. Distributions of galactic latitudes and dispersion measures for the candidates are presented there. Despite the good enough correspondence of presented distributions to those expected of the cosmic sources, we found some features, that could be better explained by the hypothesis of some number of weak interference detection. It became obvious that it is necessary to have a tool which allows to visualise and explore all the processing stages of each identified “candidate” in order to identify and eliminate much more reliably of the interference among signals which do not look like typical interference. It would be possible if we carry out a detailed study of the spectral and temporal properties of each signal at all stages of processing. For this purpose searching procedures for initial data in the data array of about 100 Tb, routines for de-dispersion, filtering and integration across the whole band and the selected partial bands have been de-

*i.p.kravtsov@gmail.com

© I. P. Kravtsov, V. V. Zakharenko, I. Y. Vasylieva, S. S. Mykhailova, O. M. Ulyanov, A. I. Shevtsova, A. O. Skoryk, P. Zarka, O. O. Konovalenko, 2016

veloped. Furthermore, a possibility of simultaneous display of the results of all stages of data processing and analysis of changes in fine-tuning of the time delay settings, filtering and integration sub-bands, etc. are provided.

Full data processing pipeline consists of two parts: fully automatised part — “long pipeline” (which detects a large amount of values (signals) exceeding selected threshold level which equals 5.5 RMS, and semi-automatised — “short pipeline”, which allows to test conformity of each individual signal to our criteria in more detail. And this, in turn, allows to weed out a huge amount of RFI that cannot be identified automatically. During the data processing we actually learn to distinguish transient signals from noise as authentically as possible, and for this purpose we selected several criteria.

One of these criteria is signal broadbandness, i. e. if the signal is narrowband — then most likely it is RFI. To check the broadbandness the aforementioned “short pipeline” is used. It allows to look at the spectrum of the signal and to analyse it by frequency. Spectrum of the useful (cosmic) signal has to be close to Gaussian, broadband, not be torn, rectangular, etc. In the “short pipeline” we change the frequency band and it helps us to separate natural and artificial signals. Remarkably, pulsar pulses are not always very broadband: if bandwidth is 16 MHz pulse width may be 3, 5 or 10 MHz. Another important fact is that “long pipeline” worked only with individual signals. It means that we always received a comparison of average background and excess of a separate individual signal over this background. We did not analyse the interference for repeatability, but now “short pipeline” is able to analyse individual and repetitive pulses separately. This allows to spot repetitive interference signals with high efficiency.

The main window of our program is shown in Fig. 1. Some parameters are tunable, including a maximum of signal intensity in relation to the calculated RMS noise over 16 adjacent data frames (1 frame = 1 horizontal strip with ~ 32 seconds duration, a complete data file consists of 16 frames), but only in the vicinity of the dispersion measure, which interests us. Smoothing parameter (`smpar`, see Fig. 1) is the option, which cuts the high frequencies off. Similarly, to improve the contrast we remove the slow variations of background (adjusted by `smpar_b`, i. e. smoothing parameter in background). We can choose a DM value and vary it with a 0.002 pc/cm^3 step and analyse individual and repeated pulses in some part of the file (or in the whole file, parts = 1, or 1/2, 1/4 of file and so on, i. e. parts = 2 and 4 respectively, to 1/16), and in each of 16 parts of file we start analyse from.

First, we can change the cut-off level of “signal-to-noise-ratio” (SNR) and then our “magnifying glass” (three-dimensional figure, which shows the SNR versus time and DM) can make a signal pattern more

contrast (Fig. 2). We can vary the smoothing parameter from 1 to 8 (a more “wooly” figure with not very large SNR). Result of applying `smpar`-parameter is presented in Fig. 3. Parameter of background alignment could also help to improve the SNR, but we can face an unwanted consequence — centring of pre-processed data as we cut more and more low-frequency harmonics (Fig. 4). In addition, in individual view mode we can adjust the time (i. e. central position, where our “magnifying glass” looks), shift in the time coordinate, in DM, and see the full frequency band and 4 sub-bands. Also we can see cleaned output signal and frequency sub-bands (which can be adjusted by parameters “Narr” and “Wide”, so we can change the width of the sub-band) (Fig. 5). For repeated pulses, we can carry out the adjustment in the following way: change the parameter of smoothing, smoothing option of the background and select the needed part of the file. It can allow to increase SNR of pulses’ sequence and to see FFT-harmonics (Fast Fourier Transform harmonics) more clearly (especially selecting the most appropriate for analysis part of the file). SNR of pulses and FFT in marked time interval (parts & N of parts) after applying the set of parameters (threshold levels, “`smpar`”, “`smpar_b`” and DM value are presented in Fig. 6. SNR and FFT for the whole file and for “natural” threshold levels are also presented in Fig. 6. We can select the part of the file without any pulses. In this case we can see that selected FFT-harmonics are practically absent there (Fig. 7.).

RESULTS

Example of the transient is presented in Fig. 8a. Initially some signals look like signals of the cosmic origin, but primary pipeline analyses data in the whole frequency range (16.5–33 MHz). Fig. 8b shows how weak individual pulses of pulsar look like. This picture shows an unknown (at first glance) weak pulse, which is actually a pulse of PSRB1508+55. New pipeline allows to analyse signals in narrower bands, where signal is located precisely. Due to the fine tuning of parameters we can identify the most intense parts of the spectrum and can easily exclude interference, which previously seemed like signals of cosmic origin. Thanks to the relatively narrow-band signal, its curvature was much more difficult to identify than in the previous figure. Due to the tuning of parameters in the detailed analysis signs of “smearing” the maximum DM were found and appropriate signal was detected in primary data (cleaned, but not de-dispersed). After contrast adjustment, the image displays a signal not typical for natural sources, but an artificial (man-made) one. Thus, new detailed pipeline allows to distinguish RFI from the background of cosmic signals, which cannot be done by primary pipeline (Fig. 9). Example of a sub-band frequency tuning is presented in Fig. 10.

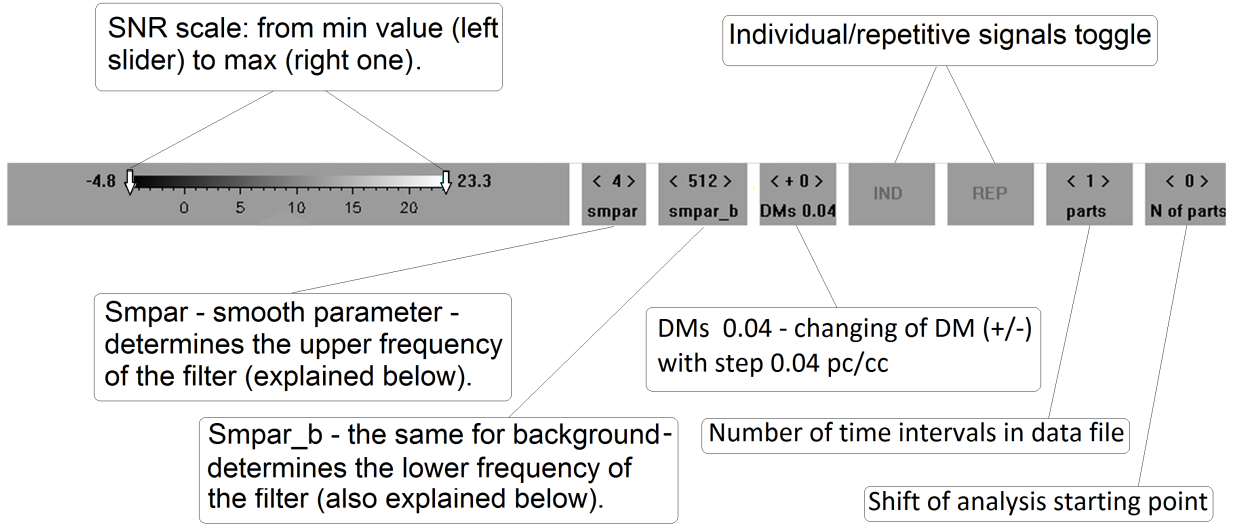


Fig. 1: The main window of the routine and tunable parameters: maximum and minimum SNR, high-pass and low-pass filters (`smpar` and `smpar_b` respectively), DM, mode (IND or REP) and choosing part of the file.

CONCLUSIONS

The main result of this work is a large number of candidates which have been weeded out using all stages of data processing. Automatic part of pipeline found tens of thousands of signals that exceed the detection threshold level (5.5 RMS) in the data. Subsequently data was visualised and inspected visually, which allowed by means of “red tired students’ eyes” to find about 1000 candidates which are the most alike to signals of cosmic sources among all candidates. And finally, these candidates were processed semi-automatically by means of the routines of “short pipeline” and it allowed to weed out almost a half of them, which were RFI (not obvious at first glance), and their identification required detailed analysis of each candidate, fine tuning of processing parameters, and spectrum analysis. By means of “short pipeline” we have also refined SNR of transients that remained after last screening: SNR of automatically selected by “long pipeline” and by human (about 1000) candidates ranged from 5.8 to about 7, and became ranging from 5.8 to 21–22, the dispersion measures of these signals were refined too. Distributions of these signals by DM and galactic latitude were analysed by us [7]. These distributions do not show any unexpected features, and thus indicate the natural origin of the signals (of those remaining after the processing and verification of candidates with a “short

pipeline”).

ACKNOWLEDGEMENT

We would like to thank Professor Victor Romanovych Khalak from Université de Moncton and the organizing committee of 23rd Young Scientists’ Conference on Astronomy and Space Physics for providing the financial support which allowed us to present the results of this work at the conference.

REFERENCES

- [1] Hewish A., Bell S. J., Pilkington J. D. H., Scott P. F. & Collins R. A. 1968, *Nature*, 217, 709
- [2] Konovalenko A., Sodin L., Zakharenko V. et al. 2016, *Exp. Astron.*, 42, 11
- [3] Kramer M., Lyne A. G., O’Brien J. T., Jordan C. A. & Lorimer D. R. 2006, *Science*, 312, 549
- [4] McLaughlin M. A., Lyne A. G., Lorimer D. R. et al. 2006, *Nature*, 439, 817
- [5] Vasylieva I. Y., Zakharenko V. V., Konovalenko O. O. et al. 2014, *RP&RA*, 19, 197
- [6] Vasylieva I. Y., Zakharenko V. V., Zarka P. et al. 2013, *Odessa Astron. Publ.*, 26, 159
- [7] Zakharenko V. V., Kravtsov I. P., Vasylieva I. Y. et al. 2015, *Odessa Astron. Publ.*, 28, 252
- [8] Zakharenko V. V., Vasylieva I. Y., Konovalenko A. A. et al. 2013, *MNRAS*, 431, 3624

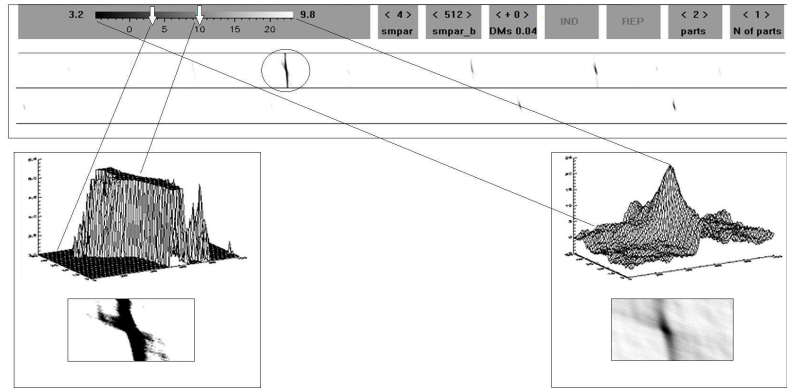


Fig. 2: Tuning the cut-off levels of SNR. The figure shows setting of a new maximum and minimum SNR values and reprocessed data.

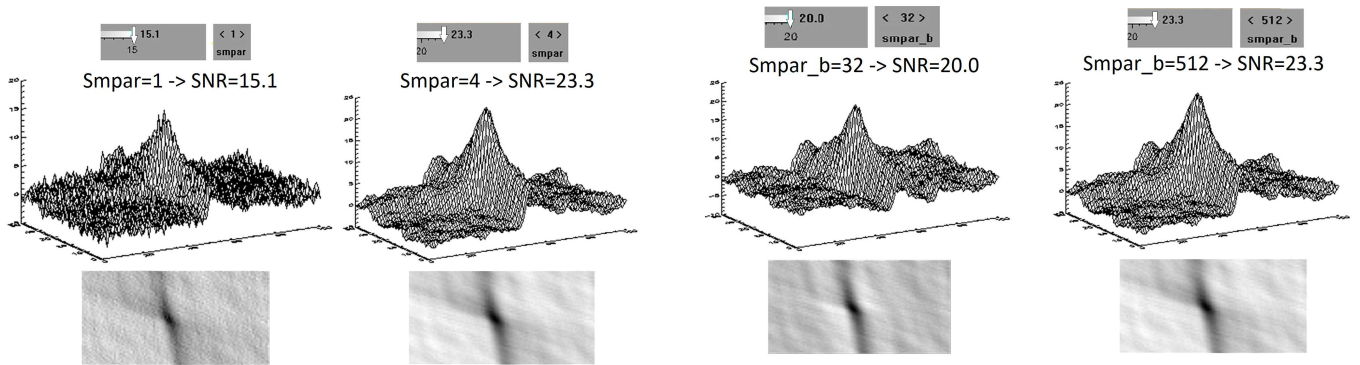


Fig. 3: Tuning the low-pass filter cut-off (smpar-parameter applying).

Fig. 4: Result of smpar_b-parameter applying which cuts low frequency components (slow background variations).

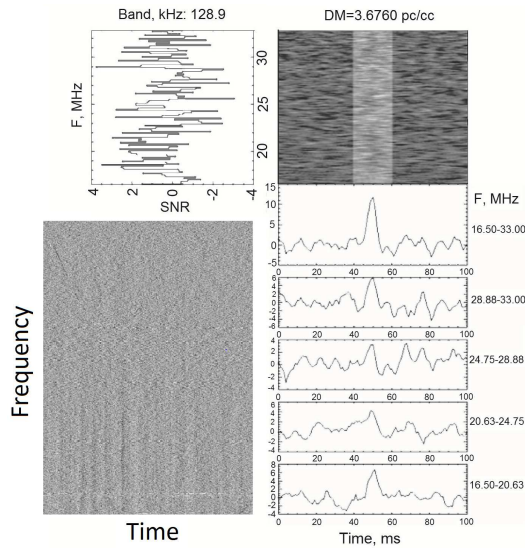


Fig. 5: Individual pulse analysis. Interface includes the following panels: dispersed pulse, stacked spectrum of the de-dispersed pulse by time and summation over frequencies.

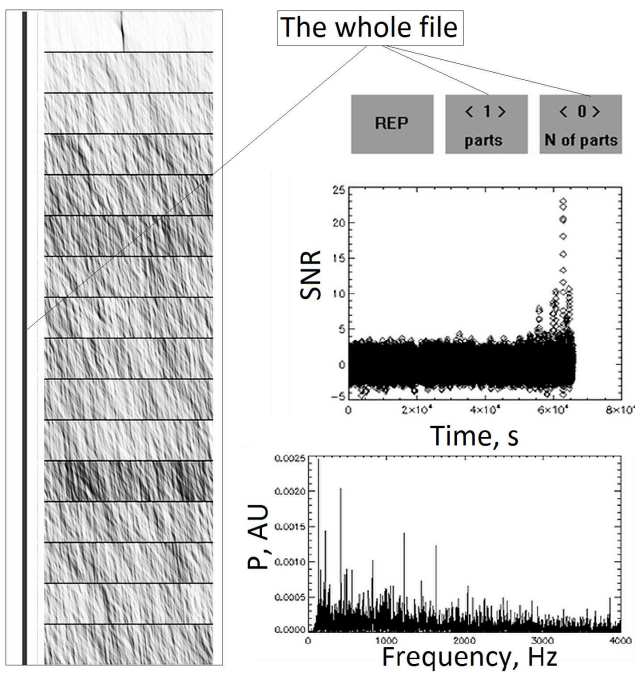


Fig. 6: Repeated pulse analysis. The panels show both a time sequence and its FFT. They can illustrate maximal outbursts and pulse harmonics.

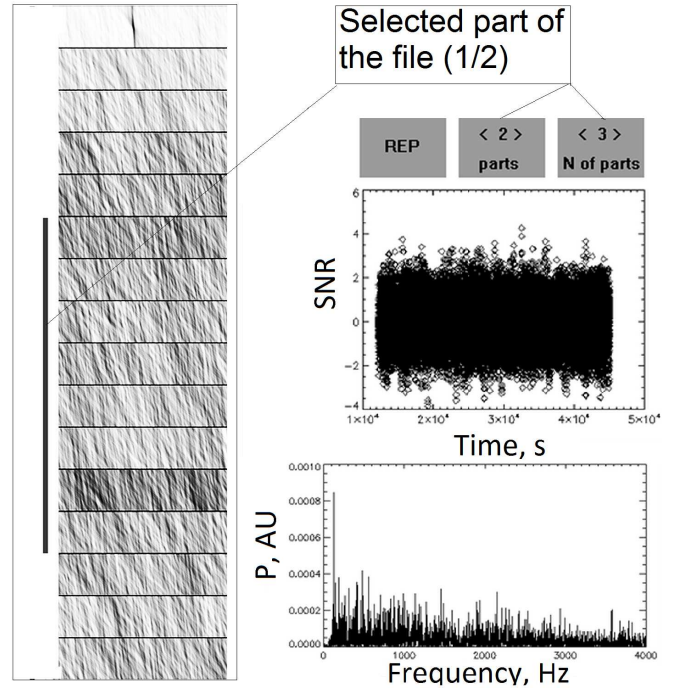


Fig. 7: Recalculated SNR and FFT panels of selected part of the file.

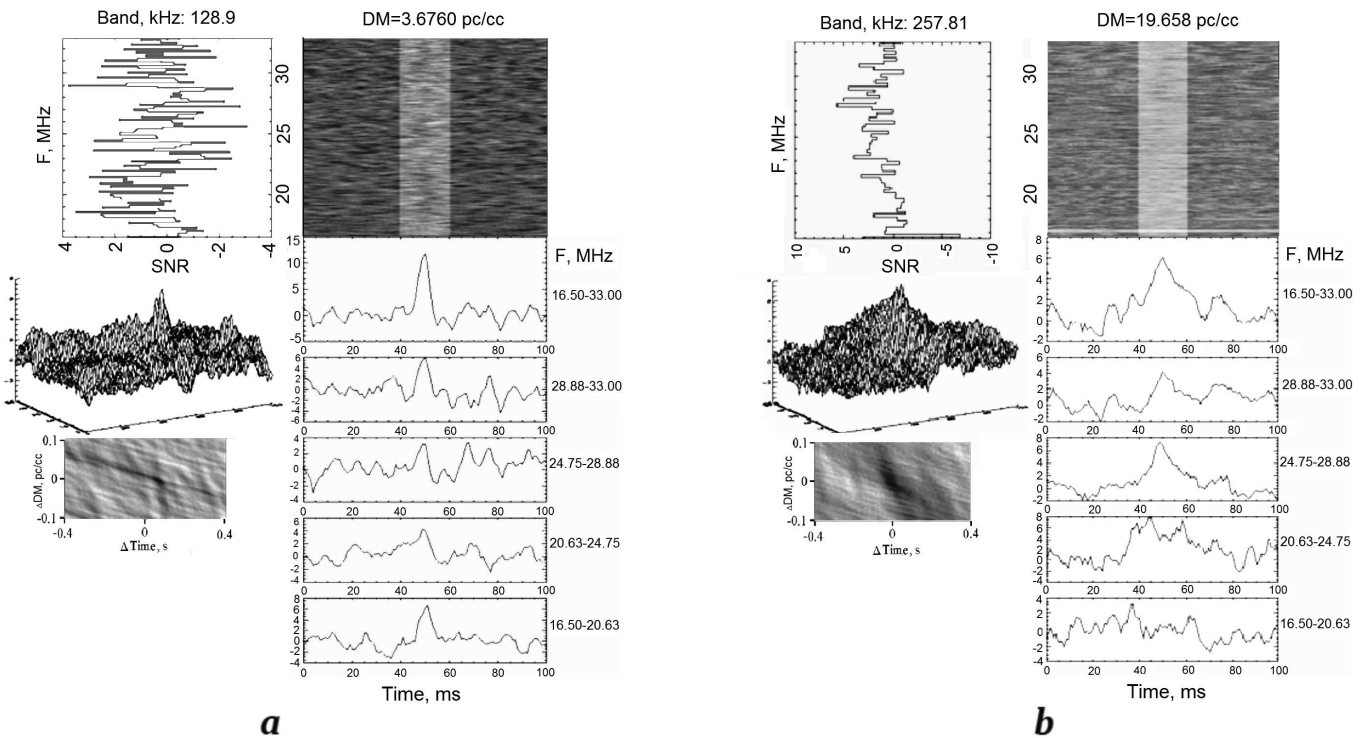


Fig. 8: *a*) Example of a transient signal from an unknown source. *b*) Weak pulse of PSRB1508+55.

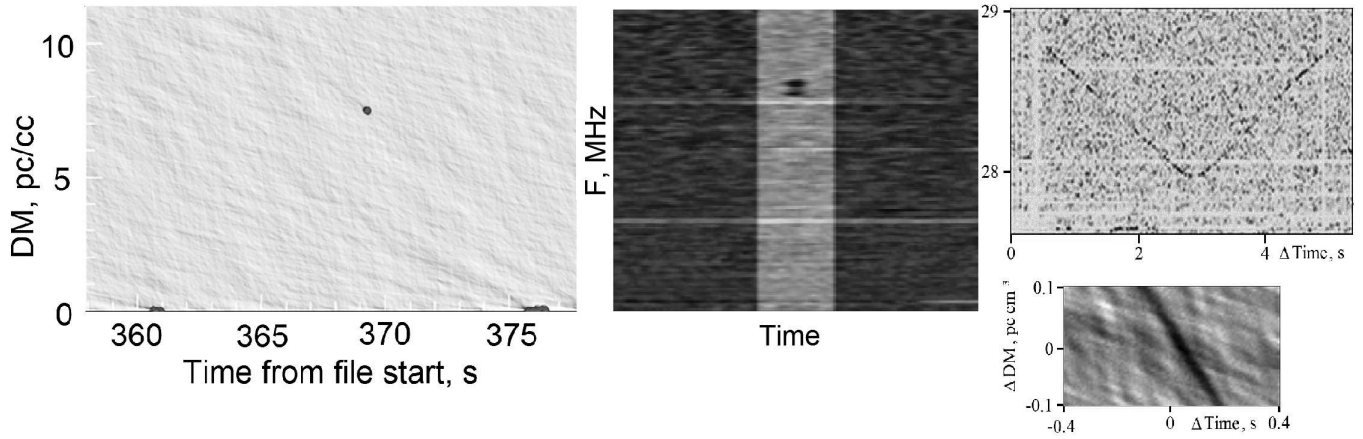


Fig. 9: Example of RFI found by means of "short pipeline". Left panel shows the signal found by means of automatic routine ("long pipeline"). Middle and right ones - results of the verification routine work ("short pipeline").

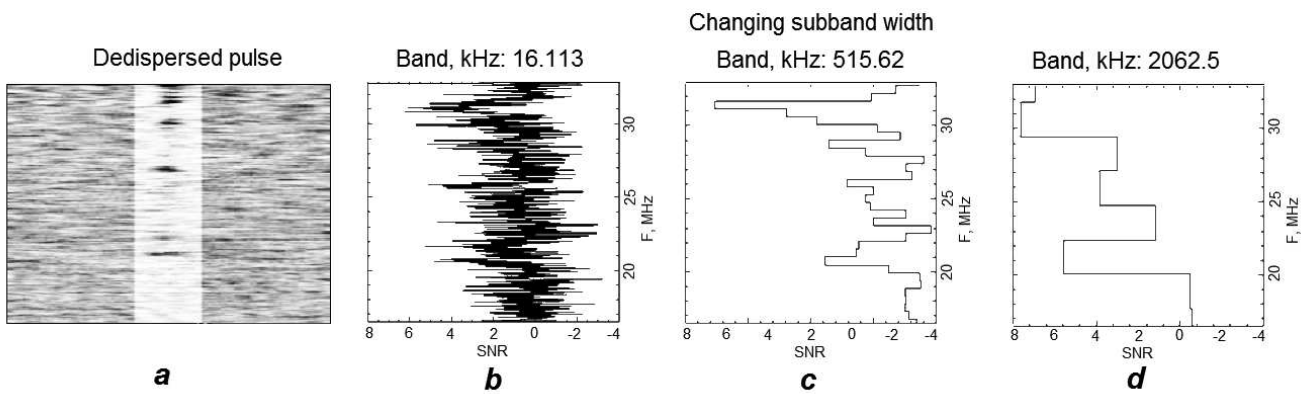


Fig. 10: De-dispersed spectrogram of pulse (a) and changing the width of the frequency sub-band: (b) 16 kHz; (c) 516 kHz; (d) 2062 kHz.