Classification of diagnostic features of transient signals in the electric power industry

Rustem Khuziashev[1,*], *Igor* Kuzmin[1], *and Iluza* Irkagalieva[1]

[1]Kazan State Power Engineering University, 420066 Kazan, Russian Federation

Abstract. Problems of practical implementation of traveling wave fault location caused by the registration of signals of different nature are considered. Analysis of the experimentally recorded traveling waves made it possible to divide them into 4 groups caused by partial discharges, lightning overvoltages, scheduled switching and fault commutations. The network dispatchers only needs the fault commutation information. Traveling waves recorded near the place of their origin have different meanings of diagnostic signs. The magnitude of the pre-alarm noise, the number of pulses in the signal and the duration of the signal are used as diagnostic indicators. These three diagnostic signs allow one to recognize each of the 4 causes of the travelling waves.

1 Introduction

The intensive practical development of wave methods for determining the location of damage in the last decade is due to advances in electronic technology and communications. Despite the results of a large number [1-4] of theoretical and experimental studies of travelling wave methods of fault location (TWFL), its large-scale implementation is complicated by many problems. They are manifested both in the difficulty of achieving the potential accuracy of TWFL, and with problems related to the identification of transient signal (TS) with their causes. This implies a visual assessment of the TS and makes it difficult to automate the TWFL. If the first problem is closely related to the influence of the dispersion mechanism, which aggravates the situation especially in heterogeneous cable and air lines, the second problem complicates its implementation, when information about non-emergency TS distracts grid dispatchers from operational work.

2 Problems of introducing wave determination of the fault location

Registration of TS with a high sampling rate of up to 10 MHz delivers a large amount of data. Their intelligent processing allows you to determine not only the location of the fault, but also the cause of the TS. These include partial discharges, short circuits, and scheduled switching. The development of intelligent data processing methods will allow one to quickly deliver to the dispatcher only information about emergency events, and, after appropriate processing, reports on the state of isolation in different parts of the network. Work in this direction is still in its infancy. This is due to the complexity of developing both hardware and algorithmic

support for TWFL. The development of these works is an urgent direction, as it allows us to significantly improve the quality of information about the state of the most important object of the electric power industry-cable and overhead lines.

The TS is a time-limited radio signal or a high-frequency bearer. It is superimposed with amplitude, phase, and frequency modulations. All of these modulation types provide useful information about the cause and source of the signal generation and the parameters of the discontinuity along the signal propagation path. Finding the relationship between the parameters of the network in which the TS is distributed and the parameters of the TS itself is a difficult task.

This relationship should be found by modeling, for example, in the PSCAD software package. This is the content of the first approach. The second approach to finding this connection is based on the analysis of experimental TS, supplemented with information from the services operating the network about the causes of their occurrence, meteorological and thunderstorm information. The TWFL uses only one, the most important parameter of the TS – the global scale time of its registration at a specific point in the network. To answer the question about the reasons for generating the TS, you must use the other TS parameters. Intelligent analysis of the waveforms of the TS involves finding a connection between its structure and information about the location and causes of the generating TS.

An example of the first approach is [5]. Based on modeling in the PSCAD software package, it is shown that the TS that came to the observation point consists of a set of single video pulses due to multiple re-reflections from the ends of the line and places of heterogeneity. The time intervals between pulses contain information about the location of the TS. Proposed to determine the

---

* Corresponding author: 142892@mail.ru

place of its occurrence by finding the maximum correlation coefficient between the recorded pulse sequence and the reference sequences for all possible points of occurrence of the TS. Unfortunately, in experimental measurements, probably due to a much larger number of inhomogeneities in the real line, for example, in the form of a traverse of power transmission poles, it is impossible to distinguish elementary pulses – they all merge with each other.

The disadvantage of the approach to signal feature detection described above is that it is based on understanding the nature of the signal occurrence, which is not always available to the researcher. The problem of signal classification based on the analysis of their formal features expressed in terms of basic functions is considered in [6]. This approach is attractive because the obtained features are not related to the physical model of the signal, i.e. it is universal. The criteria for attribute quality are strongly related to the methods of filtering features, which are already known to a large number at the moment. An overview of feature filtering methods can be found in [7].

Many publications are devoted to the classification of electrocardiogram signals. It is necessary to note a close analogy in the mechanism of TS formation in power lines and the processes of pressure changes in the arteries of a living organism caused by the opening (switching) of the heart valve. A significant difference between both signals is shown in the fact that the TS is a radio signal, and the QRs signal of the electrocardiogram is a video signal, due to the absence of reflections from the end inhomogeneities in the blood arteries, which smoothly pass into the capillaries. Reference [7] provides an overview of various numerical methods for processing electrocardiograms for detecting the QRS signal.

TS generated by emergency switching have a large amplitude and are easily detected, but as they propagate, their amplitude decreases significantly, and the structure is greatly distorted. Of course, this makes it difficult to apply mathematical methods for detecting TS [8] and forces us to use algorithms that take into account the physics of their occurrence [9].

Neural networks allow one to work with unstructured information, and they themselves extract the necessary features for classification, saving the researcher from the need to determine the features [10,11]. The disadvantage of neural networks is their training procedure, which makes it difficult to use them in the analysis of TS.

The proposed work looks for the connection of the cause of the TS with its parameters based on the second approach, based on visualization of experimental data, and the development of a heuristic algorithm.The aim of feature selection is to determine a feature subset as small as possible. It is the essential preprocessing step prior to applying data mining tasks. It selects the subset of original features, without any loss of useful information. It removes irrelevant and redundant features for reducing data dimensionality.

3 Basic types of transient signals

In April 2019, on the cable line 110 kV 11.058 km long in Kazan electrical grids, a software and hardware complex of traveling wave fault location was installed as part of sensor No. 23 installed at substations (SS) Vostochnaya, and No. 29 installed at the Centralnaya Substation (Fig. 1). Both sensors register transient signals generated by commutations of any nature using split current transformers connected to the secondary circuits of standard measuring current transformers of two phases of the cable line "A" and "C" [4]. Reference [4] also consideres the algorithm of operation with a larger number of sensors.

During the operation, seven synchronous events were recorded simultaneously at both substations and many of single events recorded only on each side of the line. Acquiring waveforms with a short sampling interval of 1.075 microseconds on a global time scale allows you to determine the location of the TS.
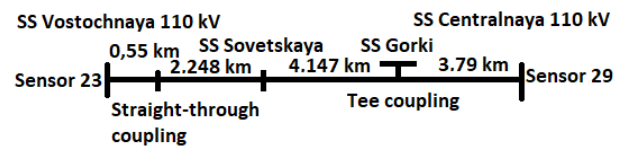
Fig. 1. Cable line diagram. Substation Vostochnaya - Substation Centralnaya.

3.1 Algorithms for determining the essential features of the transient signals

Consider the sequence of determining the significant parameters. The oscillogram registered in global time scale and transmitted to the upper processing level consists of 5652 samples, which corresponds to total duration of more than 6ms. The oscillogram is registered by blocks, which allows analyzing the pre-accident part of the TS waveform. It contains important information about the value of the constant signal level and the spread of the amplitudes of instantaneous samples relative to the value of the constant signal level. The beginning of the transient signal occurs in the interval of sample numbers from 1500 to 2500. For the analysis of pre-accident information, we selected samples from 1 to 1000 with a total duration of more than 1 ms. At this interval, the average of the instantaneous amplitudes of the samples (AVG) is calculated. Then all the instantaneous amplitudes of the waveform are shifted by the above calculated value.

The displaced amplitude values are used to calculate the value of the maximum deviation of the instantaneous amplitude (MAX) and the standard deviation (RMSD) in the pre-emergency time interval. They both together are a measure of the dispersion of the noise deviations of the instantaneous amplitudes. These values are used as a threshold value for determining the onset of a transient signal (OTS) from displaced instantaneous amplitudes. The presence of a TS in the oscillogram is determined by the excess of the instantaneous amplitude of the oscillogram over the threshold in the time interval from

1001 counts to the end of the oscillogram. The time (MTS) of the instantaneous amplitude maximum deviation of the transient signal in the interval after the OTS and its value (VMTS) are determined. The duration of the leading edge (TLE) is calculated as the time interval between the position of the MTS and the OTS.

The envelopes of the TS from above and from below are determined in the form of step functions passing through the local maxima and minima of the displaced values of the TS. The swing is determined as the difference between the TS envelopes from above and from below. At the step function of the swing range, the heights of each step are determined and the displacement of the temporary position of the maximum increase in the swing range relative to the OTS (DOTSP) is determined.

The specific, per discrete time interval, energy (UE) is calculated as the sum of the squares of the module of the instantaneous displaced amplitudes of the oscillogram exceeding the MAX, divided by their number. The normalized duration of the TS (DTS) is calculated as the result of dividing the UE value multiplied by the number of oscillogram samples by the square of the MTS value.

To calculate the number of pulses in the TS, a threshold value is set in percentage (P%) of the VMTS, which is compared with the modulus of the amplitude of each instantaneous sample of the displaced amplitude values. The result of the comparison is entered into the Z_i series, each member of which takes zero or one value depending on the result of the comparison with the threshold. For the time beginning of the next impulse (Tn1, Tn2, Tn3, Tn4), the temporal position of the first unit term of the series Z_i is taken, if the time duration of the zero terms (ZTD) of the series Z_i is located before it. For the end of the next impulse (Tk1, Tk2, Tk3, Tk4), the time position of the last unit term of the series Z_i is taken, if the ZTD of the zero terms of the series Z_i is located after it.

To calculate the half-period of free oscillations (FO) of the TS, a series T_i is formed, each term of which is the result (-1 or 1) of comparison with the zero of the instantaneous amplitude of each sample of TS displaced amplitude. From the T_i series, the $D_i = T_i - T_{(i-1)}$ series is formed. From the series D_i, a series of the current sum of the half-period samples is formed. The average value of the value of the half-period of the first TS pulse is calculated as the arithmetic mean of the nonzero values of the D_i series in the time interval from Tn1 to Tk1.

3.1 Algorithms for determining the causes of the transient signals

Figures 2-5 shows the feature oscillograms of the TS recorded on both sides of the line. During scheduled switching, the TS consists of several pulses each with a duration of more than 100 microseconds, due to the contact bounce of the high-voltage switch. These pulses, when propagated over the line, due to the dispersion mechanism, merge and are recorded as a single pulse with a free oscillation frequency of less than 10 kHz. (Fig.2).

In thunderstorm overvoltages, the TS, due to the mechanism of streamer formation, is preceded by increasing noise, and its main body is created by the main lightning discharge. Only the main body of the TS is registered (Fig.3), when this TS propagated along the line on the big distance.

TS caused by partial discharges (PD) is characterized by a short duration (less than 50 microseconds) and a high frequency of free oscillations (more than 100 kHz) (Fig.4).
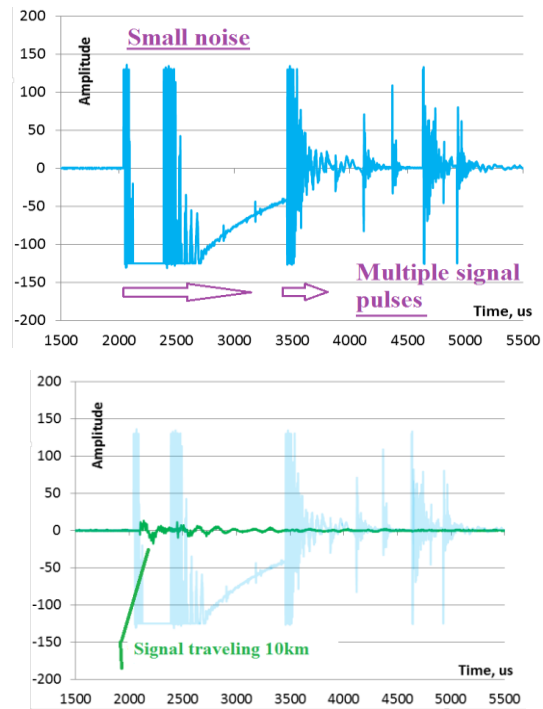

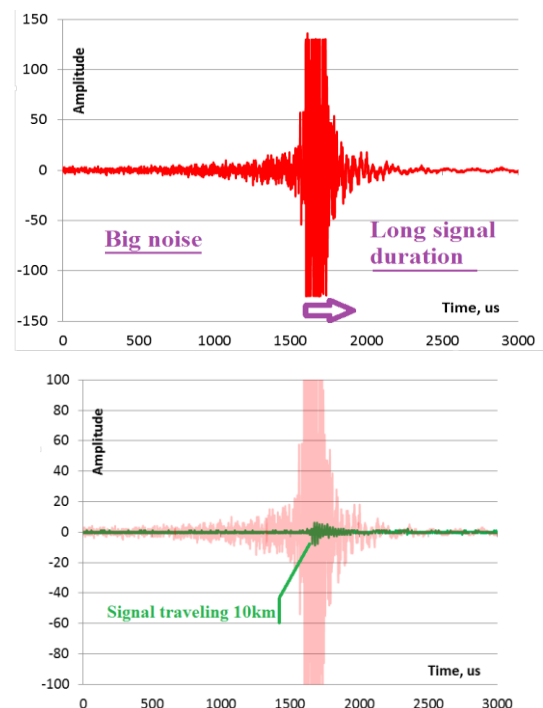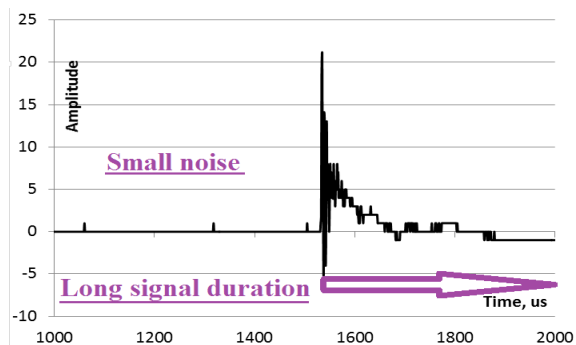**Fig. 2.** Typical example of scheduled switching TS.
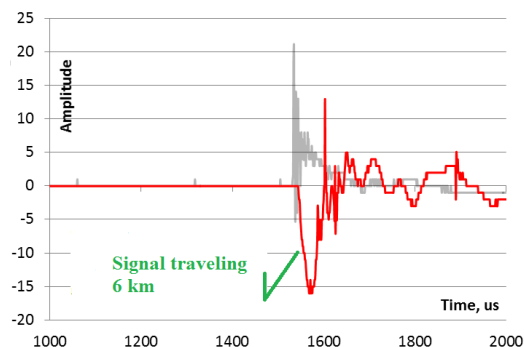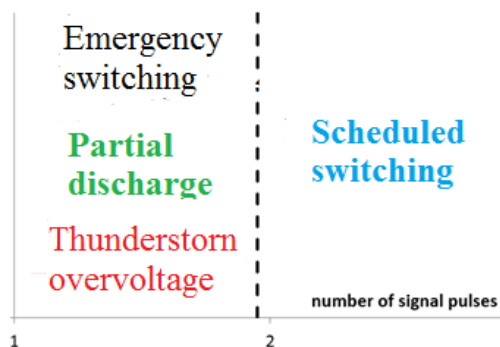

Fig. 3. Typical example of thunderstorn overvoltage TS.
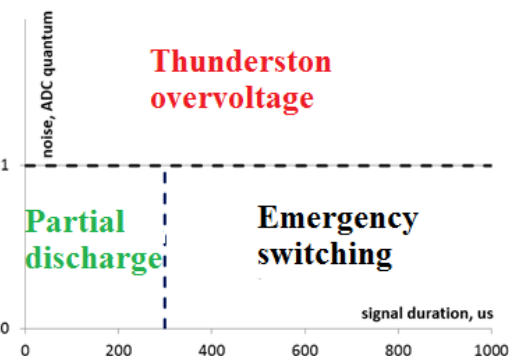
Fig. 4. Typical example of single-phase earth fault TS.
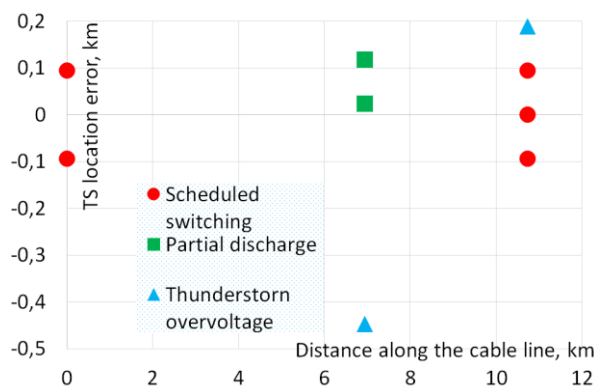

Fig. 5. Typical example of single-phase earth fault TS.


Fig. 6. Division by number of pulses.

Fig. 7. Division by noise and signal duration.

**Fig. 8.** Location error of the TS source.

4 Statistics of synchronous and single events

Three years of operation of the software and hardware complex for wave determination of the fault location in Naberezhno-Chelny RES (Fig. 9) made it possible to obtain a large volume of oscillograms of signals of transient processes. Many synchronous operations of the sensors of the complex were recorded (Fig. 10, 11, 12), which make it possible to determine the place of occurrence of the transient signals. The distribution by months (Fig. 10) indicates many of registered synchronous cases in the spring-summer-autumn seasons, caused by an increase in electrical discharge activity (Fig. 4) in the high-voltage line insulation due to high humidity and induced by lightning overvoltages (Fig. 3). These events are grouped by specific weather-related days. Events are grouped to specific areas of the power line that are characterized by weakened high-voltage insulation or induced lightning surges.
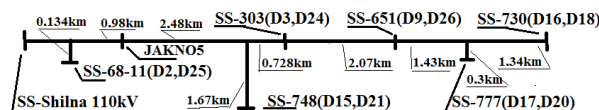

**Fig. 9.** Topology of the location of sensors of the complex in the NCHRES.

Most of these events have no effect on line performance. Some events lead to emergency damages resulting in line shutdowns by relay protection. Such an event occurred on April 10, 2018 (Fig. 13). Six short-term single-phase earth faults (SPEF) ended with phase-to-phase short circuit occurred within 55 minutes. All of them were registered by the sensors of the complex. Further, two events of supplying voltage to the line (scheduled switching) in JAKNO5 were recorded when searching for a fault location. Stable values of different time delays (Fig. 13 and Table 1) of emergency and scheduled commutations indicate different places of TSs occurrence.
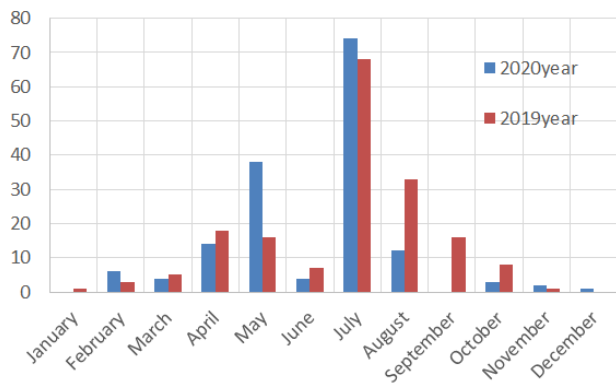
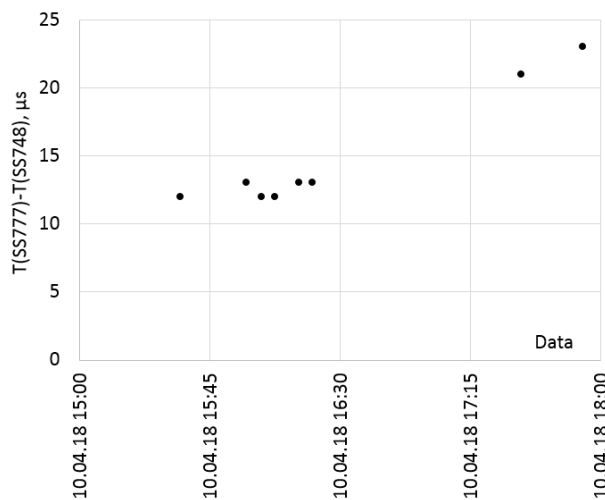**Fig. 10.** Distribution by months of synchronously registered events.
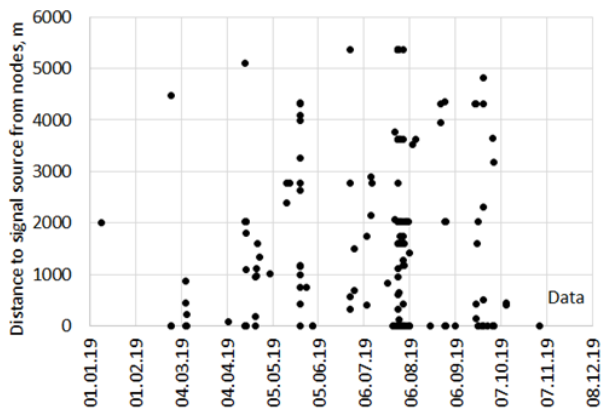
Fig. 11. Distribution by months of programmed distances.
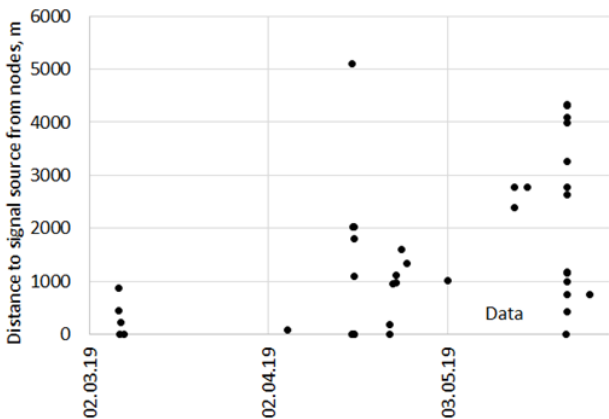
Fig. 12. Distribution by days of programmed distances.

The practical implementation of traveling wave fault location forces us to look for a connection between the cause of the transient signal and its features. Based on the analysis of experimentally recorded waveforms, numerical values of the transient signal features are derived. They made it possible to determine the cause of their occurrence: partial discharges, lightning overvoltages, scheduled switching and fault commutations. The studied essential features of transient signals have a clear physical meaning and are easily determined, which eliminates the need to use more complex methods of artificial intelligence. The result diagrams visually illustrate the desired relationship and are easily algorithmized programmatically. The results

obtained can be used in the operation of fault localization complexes. The results are initial. Subsequent studies will refine the reliability of the transient signal classification.

Fig. 13. Distribution of time delays of an emergency event.

Table 1. Time parameters of TSs registered at SPEF on April 10, 2018.

	15:34:47	15:57:48	16:03:06	16:07:34	16:15:52	16:20:26	17:32:37	17:54:01
	dT, μs	dT, μs	dT, μs	dT, μs	dT, μs	dT, μs	dT, μs	dT, μs
SS748	0	0	0	0	0	0	0	0
SS777	12	13	12	12	13	13	21	21
	Tlead, μs	Tlead, μs	Tlead, μs	Tlead, μs	Tlead, μs	Tlead, μs	Tlead, μs	Tlead, μs
SS748	1	1	1	1	1	1	1	1
SS777	21	21	20	21	20	23	37	37

5 Analytics of diagnostic features

The main task of the complex is to promptly inform the grid dispatcher about the location of the fault event. An auxiliary task of the complex is the preparation of the final reporting on the dynamics of the place of origin and the intensity of electric discharge activity. To solve both problems, it is necessary to solve the problem of classifying the type of registered TSs according to the features in Fig. 2-7. The classification problem is solved based on the analysis of the set of diagnostic features [12].

Diagnostic features include the parameters of the TS, such as the start time, maximum amplitude, number of pulses in the signal, the maximum amplitude of noise, etc. To study the information content of various diagnostic features, a program was compiled that calculates these features in the recorded TSs. The calculated diagnostic features form a database that can be visually analyzed by standard analytical methods - correlation dependences of pairs of features. The diagnostic signs of all TSs (single and synchronous)

were processed (Table 2). The increased number of PDs on SS with a voltage of 110 kV is explained by the large number of PDs on the higher voltage line. The increased number of PDs caused by SS-730 and SS-777 indicates the existence of high voltage insulation problems on these SSs. The number of registrations of TSs are varied throughout the year (Fig. 10).

Table 2. The number of TSs registered from 10.30.2021 to 11.21.2021.

SS	6811	748	303	651	777	730	Vost.	Centr
Number	2	4	26	21	39	39	81	48

The largest number of singly recorded events are generated by PDs of different intensities. They are manifested in different forms of TSs: single short durations, a sequence of pulse signals or as continuous noise sequence. Three algorithms are used to determine the beginning of any TS. Differential algorithm "SensorFlag" is based on exceeding the programmable threshold by the difference of two instantaneous signal amplitudes shifted by 100 μs in the interval of three consecutive samples. This algorithm is implemented in the sensors of the complex and is used to store pre-emergency instantaneous amplitudes in the interval of the first 1000 μs of the oscillogram. The "FLAG" algorithm implemented on the server of the complex searches for the time position of the first instantaneous amplitude, earlier than the time of the maximum signal value, which exceeded the threshold equal to three standard deviations of the instantaneous amplitudes of the first 1000 μs of the oscillogram. The "mcf2StartPointA" algorithm implemented on the server of the complex searches for the time position of the first instantaneous amplitude that has exceeded the threshold equal to the maximum instantaneous amplitude for the first 1000 μs of the oscillogram. Figure 14 shows the correlation dependence of the FLAG and SensorFlag features. Events located below the linear dependence correspond to PD with duration less than 3 μs, not recorded by the "SensorFlag" algorithm. Events located above the linear dependence correspond to a sequence of PD pulses with increasing amplitude. In Fig. 15, the events located below the linear dependence also correspond to PD with a duration of less than 3 μs, not recorded by the "SensorFlag" algorithm. Events located above the linear dependence correspond to a sequence of PD pulses of large amplitude with a duration of less than 3 μs recorded in the interval of the first 1000 μs of the oscillogram.
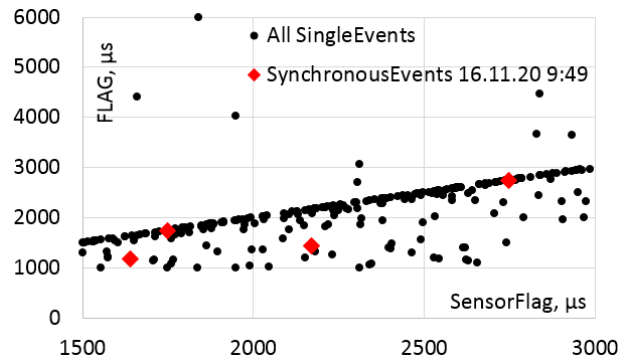

**Fig. 14.** Correlation dependences of the times of the beginning of the TSs.
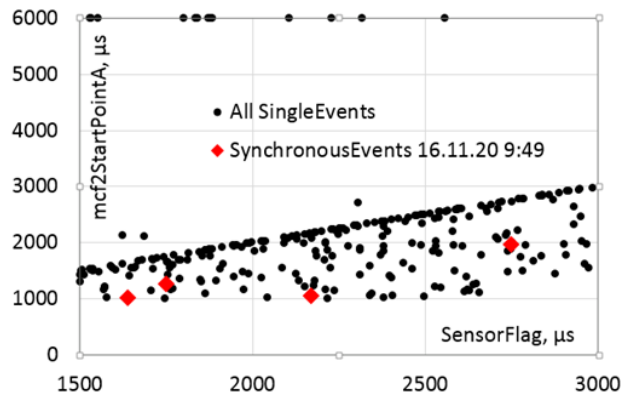
Fig. 15. Correlation dependences of the times of the beginning of the TSs.

The highlighted diagnostic signs carry important information about the parameters of the TS, the reason for their occurrence and contain information about the state of high-voltage insulation at different points of the network. Let's analyze our database for 11.16.2020 of the diagnostic features of the TSs obtained on the section of the 10 kV hybrid overhead cable network of Naberezhno-Chelninskie electric grid shown in Fig. 9. The onset of electrical discharge activity was recorded in three phases of SS-777 (Fig. 16). After 8 ms, the increased intensity of the electric discharge activity was synchronously recorded on SS-651 and SS-730 (Fig. 17,18). The values of the duration of the TS and the duration of its leading edge make it possible to determine the remoteness of the origin of the TS from the place of its registration.

According to the oscillograms of the TS on the SS-777 buses, the PD intensity in phase B is significantly higher. Phase B signal sequence started early by about 800μs. The period of free oscillations of the TS at all phases varies in the range from 3 to 6 μs, which indicates the close location of the PD source.
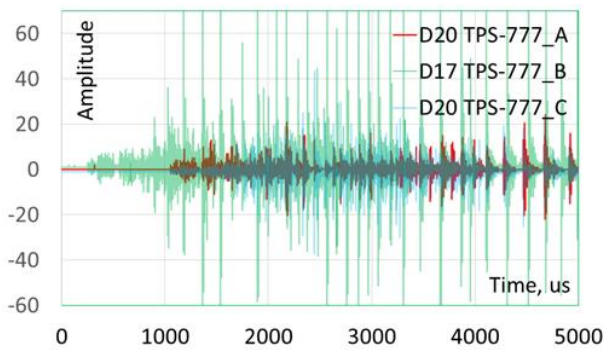
Fig. 16. TS on SS-777 11.16.2020 at 09:49:54 0 ms.
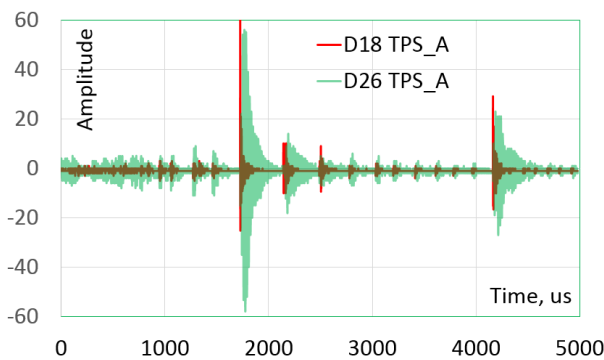
Fig. 17. TS on SS-730 and SS-651 11.16.2020 at 09:49:54 8 ms.
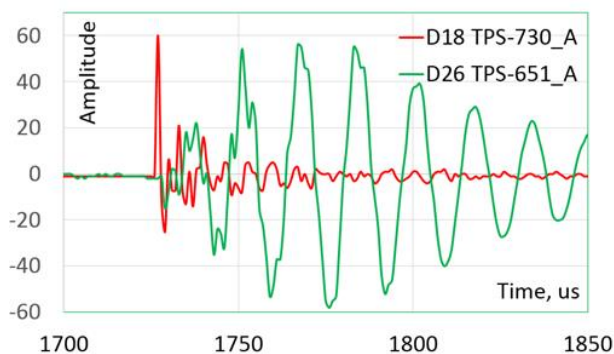
Fig. 18. TSs on SS-730 and SS-651 11.16.2020 at 09:49:54 8 ms.

As can be seen from Fig. 18, the beginning of the TS in SS-651 was recorded 2 μs later than in SS-730, which is consistent with different lengths of the hybrid line from the tap-off feeder node at SS-777 (Fig. 9). A characteristic feature is the change in the shape of the TS with an increase in the distance traveled. The duration from the beginning of the TS to its maximum amplitude and the period of free oscillations increase. Diagnostic features in Fig. 14 and Fig. 15 allow to select the sequence pulses from which the above synchronous events consist. This allows the event to be classified as PD.

6 Conclusions

The practical implementation of traveling wave fault location forces us to look for a connection between the cause of the transient signal and its features. Based on the analysis of experimentally recorded waveforms, numerical values of the transient signal features are derived. They made it possible to determine the cause of their occurrence: partial discharges, lightning overvoltages, scheduled switching and fault commutations. The studied essential features of transient signals have a clear physical meaning and are easily determined, which eliminates the need to use more complex methods of artificial intelligence. The result diagrams visually illustrate the desired relationship and are easily algorithmized programmatically. The results obtained can be used in the operation of fault localization complexes. The results are initial. Subsequent studies will refine the reliability of the transient signal classification.

References

1. L. Hongyu, U. Yanlei, L. Le, Y. Pengfei, Traveling wave fault location method analysis and prospect, 2020 5th Asia Conference on Power and Electrical Engineering (ACPEE), 1419-1423 (June 2020)

2. L. Xun, L. Shun-gui, H. Rong-hui, A. Jingwen, A. Yunzhu, C. Ping, X. Zhengxiang, Study on accuracy traveling wave fault location method of overhead line – Cable hybrid line and itsinfluencing factors, 2017 Chinese Automation Congress (CAC), 4593-4597 (October 2017)

3. J. Huibin, An Improved Traveling-Wave-Based Fault Location Method with Compensating the Dispersion Effect of Traveling Wave in Wavelet Domain, Mathematical Problems in Engineering, **1** (2017)

4. R.G. Khuziashev, I.L. Kuzmin, V.D. Vasiliev, S.M. Tukaev, Practical implementation of the wave method for determining the location of damage in branched distribution electrical networks 6 (10) kV, Electricity, Transmission and distribution, 98-107 (2019)

5. A.L. Kulikov, A.A. Loskutov, P.S. Pelevin, Algorithm for identification of a damaged section on cable and overhead power lines based on recognition of wave portraits, Electricity, **3**, 11-17 (2018)

6. V. Viculin, Automatic Feature Extraction for Signals Classification, XIX International Conference of Data Analytics and Management in Data Intensive Domains, 15-20 (October 2017)

7. K. Sutha, J. Tamilselvi, A review of feature selection algorithms for data mining techniques, International Journal on Computer Science and Engineering, **7 (6)**, 63-67 (2015)

8. B.U. Kohler, C. Hennig, R. Orglmeister, The principles of software qrs detection, IEEE Engineering in Medicine and Biology Magazine, **21 (1)**, 42-57 (January 2002)

9. C. Abadie, T. Billard, T. Lebey, Numerical signal processing methods for partialdischarge detection in more electrical aircraft, 2016 IEEE International Conference on Dielectrics (ICD), **1**, 540-543 (July 2016)

10. N.T.H. Anh, T.H. Hoang, D.T. Dung, V.T. Thang, T.T.Q. Bui, An artificial neural network approach for electroencephalographic signal classification towards brain-computer interface implementation, 2016 IEEE RIVF International Conference on Computing Communication Technologies, Research, Innovation, and Vision for the Future (RIVF), 205-210 (2016)

11. T. Debnath, M.M. Hasan, T. Biswas, Analysis of ecg signal and classification of heart abnormalities using artificial neural network, 2016 9th International Conference on Electrical and Computer Engineering (ICECE), 353-356 (2016)

12. You-Jin Park, Shu-Kai S. Fan, Chia-Yu Hsu, A Review on Fault Detection and Process Diagnostics in Industrial Processe, Processes, **8 (9)**, 1123 (9 September 2020) DOI: 10.3390/pr8091123