Development of precursors recognition methods in vector signals

V G Kapralov¹³, V V Elagin¹, E G Kaveeva¹, L A Stankevich¹, M M Dremin², S V Krylov², A E Borovov¹, H A Harfush¹ and K S Sedov¹

¹ SPbPU, Saint-Petersburg, RF
² NRC «Kurchatov Institute», Moscow, RF
³ E-mail: kapralov@phtf.stu.neva.ru

Abstract. Precursor recognition methods in vector signals of plasma diagnostics are presented. Their requirements and possible options for their development are considered. In particular, the variants of using symbolic regression for building a plasma disruption prediction system are discussed. The initial data preparation using correlation analysis and symbolic regression is discussed. Special attention is paid to the possibility of using algorithms in real time.

1. Introduction

As the design of plasma machines becomes more complex and the understanding of processes that occur in them becomes more profound, and due to the increase in productivity of computing systems and the performance of measuring and actuating devices, as well as development of new mathematical methods and algorithms, the task of predicting the behavior of plasma and predicting any threshold events in plasma discharge arises more often, including the conditions of processing multichannel data in real time.

Thus, a typical problem for tokamaks is the prediction of plasma disruption. Plasma disruption is a transient event during which the plasma discharge ceases. During the plasma quench, all of the energy that is stored in it during the discharge time leaves the plasma cord, which leads to large loads on the structural elements of the tokamak and the surface of the first wall of the vacuum chamber. In addition, during the plasma quench, runaway electron beams can be formed, which have entered into the continuous acceleration conditions. These beams, upon reaching the chamber wall, can locally melt its surface. To reduce the negative effects of plasma disruption, a number of methods and systems have been developed to mitigate the quench such as impulse massive gas injection. However, for the operation of such systems, up to 10 ms is needed. Thus, it is necessary to predict the time of disruption and generate a trigger signal with the required anticipation. Therefore, a disruption prediction system is used that calculates the probability of disruption after a certain period of time in the future. If a certain threshold probability value is exceeded, a signal must be generated to trigger a controlled quench mitigation system to prevent the formation and provide suppression of runaway electron beams. The longer is the time interval for which a prediction of disruption is made, the simpler is the technical implementation of the systems that are involved in the control of the quench, and the wider range of possible systems that can be used to neutralize the undesirable consequences of the predicted disruption.

Based on the previous results on this subject [1], for ITER (using database IGDBTH4V6B), symbolic regression methods and genetic algorithms are used to form the scaling of temperature
threshold for the transition into improved confinement mode. In this mode, in a certain layer of plasma column, the transport coefficients of particles and heat decrease, which leads to the formation of a transport barrier and increases the plasma density and temperature inside it. In [2], the APODIS (Advanced Predictor Of DISruptions) system is considered, which was used on the JET tokamak for predicting plasma disruptions. A sufficient reliability of prediction with a small number of false and missed alarms is demonstrated, even for the highly noisy data. In addition, the methods for forming training data sets to detect inaccurate experimental data in real time are proposed. For APODIS, two synthetic signal generators have also been developed [3], which allow interpolation of data gaps. This allows the disruption prediction system to be operational if there are inaccurate or missed data.

The purpose of this paper was to consider approaches for taking into account during the formation of precursors not only explicit dependencies in the original data but also hidden dependencies by simply including in the number of considered parameters the time samples, which are registered separately for each of the processes with a different frequency.

2. Data selection

The formation (selection) of input data of the parameters both of the plasma during the discharge and the T-10 tokamak itself is an important stage during the construction of a system for processing vector signals, especially in real time. At this step, it is necessary to exclude interdependent signals to reduce the demand on hardware performance by discarding duplicate information at the early stage of system construction.

Primary processing was performed using a correlation analysis. As a result, a set of input signals was formed, which was an order of magnitude smaller than the set of all detected signals in the tokamak database. The primary set of parameters consisted of the total plasma current, electron temperature and density at the center of the plasma column, and the effective charge of the plasma. The correlation coefficients between these parameters are presented in Table 1.

| Parameters | Plasma current | Electron temperature | Electron density | Effective charge |
|------------|----------------|----------------------|-----------------|-----------------|
| Ip         | 1.00           | 0.16                 | -0.32           | -0.33           |
| Te         | 0.16           | 1.00                 | 0.41            | 0.21            |
| ne         | -0.32          | 0.41                 | 1.00            | 0.21            |
| Zeff       | -0.33          | -0.21                | 0.21            | 1.00            |

3. Drift of tokamak parameters

During the development of a system for predicting plasma disruptions, it is possible to consider various options of its functioning. It is desirable to consider these options during the process of system development and training.

The first such case takes into account the slow changes of machine parameters with time. The system is trained using data that was received in the past. The data are applied to signals that have not yet been registered at the time of learning. Thus, for the time since learning, the coefficients of functional dependencies could have changed somewhat, for example, due to the change in surface quality of the first wall of the vacuum chamber. Moreover, such drift of parameters can be observed within the same day of tokamak operation. The first morning shots are performed at one vacuum conditions and recycling, and later, the situation may change.

The second case that should be considered when constructing a prediction system is the possibility of taking into account the significant events that occur in the tokamak, both during the discharges and in the intervals between them. Such events include substandard shots, sudden change in vacuum
conditions, and switching on of active systems such as additional heating or injection. In this case, it is necessary to switch the disruption prediction system to pre-configured settings for the new operating mode. Preferably, this should be performed automatically.

The third case that needs to be taken into account during the development and training of the disruption prediction system is the change in the set of available signals. In this situation, the system can be switched to either work on available signals or include the calculation of missing data on available experimental data in real time.

To take into account the described situations, we introduce a number of parameters - firstly, the temporal and phase characteristics of the plasma machine. The absolute time \( t_U \), which has been proposed to count in milliseconds since 01/01/1970 (the beginning of the UNIX era), is convenient for programming, data acquisition and automation purposes. The accounting of cyclic changes in tokamak parameters is suggested to be considered using the time from the start of the cycle \( t_i^c \), duration of the cycle \( T_i^c \), duration of the tokamak idle time before the start of the cycle \( T_i^s \), and phase of the cycle \( \theta_i^c = t_i^c / T_i^c \). Index \( i \) specifies the type of cycle: 1 - experimental campaign, 2 - working week, 3 - working day, 4 - preparation and execution of one shot, 5 - plasma discharge itself.

For these time and phase characteristics, as well as the previously selected parameters given in Table 1, a correlation analysis was performed. Most of the coefficients in the correlation matrix were close to zero values.

Symbolic regression was used to extract the functional dependencies of the selected parameters on time and phase characteristics [4]. In addition to the basic arithmetic operations, the power \( x^a \) and trigonometric \( \sin(ax + d) \) functions were considered as permissible functional dependencies of the variable \( x \).

In this case, the dependences were found, the most interesting of which are the dependence of electron temperature on the phase \( \theta_{3c} \) of the cycle that corresponds to the working day

\[
T_e = T_{e0}\left(1 + 0.04\sin\left(4\pi\theta_{3c} - \pi / 2\right)\right)
\]  

(1)

and the effective charge of plasma from the phase \( \theta_{2c} \) of the cycle that corresponds to the working week

\[
Z_{eff} = Z_{eff0}\left(1 - 0.06\theta_{2c}^2\left(1 - \theta_{2c}^2\right)\right)
\]

(2)

Here, the normalization of values of the electron temperature and effective charge is carried out according to the data of the third conditioned discharge in the corresponding cycle.

Thus, in the disruption prediction system using temporal parameters, it is available to consider the influence of other parameters that are not explicitly taken into account because of the difficulties in measuring or calculating them.

4. Modeling

To test the suggested technique for taking into account the drift of tokamak parameters, a set of shots was taken from the tokamak T-10 database, which was processed with and without the time parameters \( \theta_i^c \) that were introduced in the previous paragraph. The results of the simulation are presented in Table 2.

Four models were considered, which differed in the sets of parameters that were taken into account. A basic set of parameters, including plasma current, electron temperature and density at the center of plasma discharge, were taken into account in all four models. The effective charge, the calculation of which complicates the application of model in real time, was considered only in models 2 and 4. Finally, the time parameters of phases of the working day and week were taken into account in models 3 and 4, according to expressions (1) and (2), respectively. The data in Table 2 are averaged over a set of 8 shots, which were used as input data for every model.
Table 2. Values of the probability of disruption for different sets of parameters and intervals of disruption prediction

| Model index | Inclusion of the basic set of parameters | Inclusion of effective charge in the model | Inclusion of time parameters in the model | Time before disruption when probability exceed 70% $t_{70}$, ms | Probability of disruption at 10 ms prior to it $p_{-10}$ | Probability of disruption at 5 ms prior to it $p_{-5}$ |
|-------------|------------------------------------------|-------------------------------------------|------------------------------------------|-----------------------------------------------|---------------------------------------------|---------------------------------------------|
| 1           | $+$                                      | $-$                                       | $-$                                      | 4.2                                           | 0.44                                        | 0.56                                        |
| 2           | $+$                                      | $+$                                       | $-$                                      | 4.8                                           | 0.55                                        | 0.62                                        |
| 3           | $+$                                      | $-$                                       | $+$                                      | 4.8                                           | 0.53                                        | 0.64                                        |
| 4           | $+$                                      | $+$                                       | $+$                                      | 5.1                                           | 0.56                                        | 0.71                                        |

It can be seen in Table 2 that the addition of any of the parameters under consideration improves the disruption prediction result. However, even when all parameters are considered, the probability of 70%, which is sufficient to start the quench mitigation system, reached 5 ms before the breakdown only for model 4.

5. Conclusions

During the process of developing and training a system for disruption prediction, one must take into account the slow drift of tokamak parameters with time both for a long time and during a working day. It is also important to take into account the changes in situation after disruptions or any rare but significant events such as change in the materials of a part of the first wall. It should be possible to automatically switch the modes of operation of the prediction system when active systems, such as additional heating or pellet injection, are switched on.

The disruption prediction system should allow to adaptively change the prediction time interval, depending on the quality of input data and devices that are switched on by the trigger. In addition, there should be the possibility of removing invalid data and interpolating data gaps.

The use of correlation analysis for the selection of parameters and the method of symbolic regression for the formation of functional dependencies during the construction of a system for plasma disruption prediction have been shown. Similar data processing schemes are also used in other areas, for example, when processing EEG signals [5].

Acknowledgements

The authors are thankful to the staff of the T-10 tokamak for the provided data and support. The work was carried out with the partial support from the State Corporation ROSATOM, contract No.H.4x.241.9B.17.1011.

References

[1] Peluso E, Murari A, Gelfusa M and Gaudio P 2014 *Plasma Physics and Controlled Fusion* **56** 114001
[2] Moreno R et al 2014 *Plasma Physics and Controlled Fusion* **56** 114003
[3] Rattá G A, Vega J and Murari A 2014 *Plasma Physics and Controlled Fusion* **56** 114004
[4] Awange J L and Pal B 2016 Symbolic Regression *Geospatial Algebraic Computations* (Berlin Heidelberg: Springer-Verlag) part 1 chapter 11 pp 203-216 DOI:10.1007/978-3-319-25465-4_11
[5] Stankevich L A et al 2016 EEG pattern decoding of rhythmic individual finger imaginary movements of one hand *Human Physiology* **42** 32