EFDA–JET–CP(12)05/04

J. Vega, S. Dormido-Canto, J.M. López, A. Murari, J.M. Ramírez,
R. Moreno, M. Ruiz, D. Alves, R. Felton
and JET EFDA contributors

Results of the JET Real-Time Disruption Predictor in the ITER-Like Wall Campaigns

Results of the JET Real-Time Disruption Predictor in the ITER-Like Wall Campaigns

J. Vega[1], S. Dormido-Canto[2], J.M. López[3], A. Murari[4], J.M. Ramírez[2],
R. Moreno[1], M. Ruiz[3], D. Alves[5], R. Felton[6]
and JET EFDA contributors*

JET-EFDA, Culham Science Centre, OX14 3DB, Abingdon, UK

[1]Asociación EURATOM/CIEMAT para Fusión. Madrid. Spain
[2]Universidad Nacional de Educación a Distancia (UNED). Madrid. Spain
[3]Universidad Politécnica de Madrid. CAEND UPM-CSIC. Madrid. Spain
[4]Consorzio RFX, Associazione EURATOM/ENEA per la Fusione. Padua. Italy
[5]Associação EURATOM/IST, Instituto de Plasmas e Fusão Nuclear – Laboratório Associado, Instituto Superior Técnico, P-1049-001 Lisboa, Portugal
[6]EURATOM-CCFE Fusion Association, Culham Science Centre, OX14 3DB, Abingdon, OXON, UK
** See annex of F. Romanelli et al, "Overview of JET Results", (23rd IAEA Fusion Energy Conference, Daejon, Republic of Korea (2010)).*

Preprint of Paper to be submitted for publication in Proceedings of the
27th Symposium on Fusion Technology (SOFT), Liege, Belgium
24th September 2012 - 28th September 2012

ABSTRACT

The impact of disruptions in JET became even more important with the replacement of the previous Carbon Fiber Composite (CFC) wall with a more fragile full metal ITER-like wall (ILW). The development of robust disruption mitigation systems is crucial for JET (and also for ITER). Moreover, a reliable real-time (RT) disruption predictor is a pre-requisite to any mitigation method. The Advance Predictor Of DISruptions (APODIS) has been installed in the JET Real-Time Data Network (RTDN) for the RT recognition of disruptions. The predictor operates with the new ILW but it has been trained only with discharges belonging to campaigns with the CFC wall. 7 real-time signals are used to characterize the plasma status (disruptive or non-disruptive) at regular intervals of 1ms. After the first 3 JET ILW campaigns (991 discharges), the success rate of the predictor is 98.36% (alarms are triggered in average 426ms before the disruptions). The false alarm and missed alarm rates are 0.92% and 1.64%.

1. INTRODUCTION

Due to the complex and highly non-linear coupling of events that lead to a disruption, predictive physics-driven models of these phenomena have not been established from theoretical considerations so far. As an alternative, data-driven models allow the estimation of useful relationships among several quantities to recognize the signature of an incoming disruption.

As mentioned in the abstract, the impact of disruptions in JET is a more serious issue with the ILW [1]. This article shows the results of a disruption predictor at JET, APODIS, that has been in operation in the JET RTDN during the three initial ILW campaigns (C28-C30, between August 2011 and July 2012). The objective has been to assess its prediction capabilities (success rate, missed alarms, false alarms and prediction times) for later use in next campaigns as trigger for mitigation actions. In the above ILW campaigns, the alarm generated by APODIS has been distributed through the RTDN and recorded for off-line analysis, but it has not been used to close any feedback loop.

2. DATA-DRIVEN MODELS BASED ON CLASSIFIERS

Data-driven models are based on machine learning methods. In particular, this article is centered on the development of a binary classifier to distinguish in JET between two plasma behaviors: disruptive and non-disruptive (or safe).

In general, in a binary classification problem, given some samples of the two different classes (Fig.1), the objective of a data-driven model is to build a decision function to determine the class of new samples. The decision function is computed with a dataset of samples whose class is known. This dataset is called training set. Usually, the model reliability is tested with samples of known classes (test set) that are not used in the training process. The greater the success rate with the test set, the more reliable the classifier is.

From a mathematical point of view, the samples are represented by features of distinctive nature to discriminate the class that one sample belongs to. In general, the set of features that describe a sample **x** is called the feature vector, $x \in \mathbb{R}^m$, where m is the number of features.

Support Vector Machines (SVM) is a binary classification methodology that is used in APODIS.

SVM maps the input space of the feature vectors into a new space (feature space) in such a way that the decision function in the feature space is a linear function (Fig.1). The mapping function is called a kernel. Typical kernels to train SVM classifiers are linear ($H(\mathbf{x},\mathbf{x}') = (\mathbf{x}\cdot\mathbf{x}')$ and radial basis function (RBF) ($H(\mathbf{x},\mathbf{x}') = \exp\{-\gamma|\mathbf{x}\cdot\mathbf{x}'|^2\}$) kernels. The first ones do not require extra parameters but the second ones needs as input the γ parameter.

Due to space limitations, the mathematical formulation of SVM is not described here but can be found for example in [2]. For the purposes of the article, it is important to mention that an SVM training requires the specification of 4 inputs: the training samples with the corresponding labels, the kernel to use, a regularization parameter C [2] and the kernel parameter when needed.

3. APODIS DESCRIPTION

The APODIS project started in 2008 with the aim of developing a disruption predictor for JET [3, 4] that overcomes the issues of earlier predictors. The first APODIS version for JET performed the simulation of the real-time processing of the predictor [4]. The current APODIS version for the JET ILW campaigns differs from the first one in six aspects: real-time applicability, the amount of shots considered for training, the signals (Table 1), the signal representations, an exhaustive data pre-processing to remove discharges with outlier signals and the use of high performance computing. The signals used by APODIS to identify a disruptive behaviour are more or less the same as those used in previous predictors. However, APODIS provides an important novelty in relation to other disruption predictors. It follows a multi-tier architecture (Fig.2) in which the three first tier classifiers are combined with another classifier in the second tier.

The basic APODIS units are temporal windows of 32 ms. All signals use a sampling period of 1 ms and, therefore, there are 32 samples of each signal per window. It has been demonstrated [5] that the combination of different signal representations in the APODIS architecture allows reducing the number of signals required. Thus, two signal representations per window have been used: the standard deviation of the Fourier spectrum (removing the DC component) [3] and the mean value. The APODIS training process is fully described in [4] and it will not be detailed here. However, it is important to summarize how APODIS works during a discharge. From the time instant in which the plasma current is above a certain threshold (presently 750 kA) and every 32ms, a feature vector $\mathbf{x}$ is formed. With the 32 samples per signal, the mean value and the Fourier spectrum standard deviation are computed. Because a total of 7 signals are used, $x \in \mathbb{R}^{14}$.

After completing a 32ms time window at a time instant t, the three most recent feature vectors ($\mathbf{x}(t-64)$, $\mathbf{x}(t-32)$ and $\mathbf{x}(t)$) are used by APODIS (fig. 2) as respective inputs to M3, M2 and M1. The output of each classifier provides the status (disruptive/non-disruptive) of its input feature vector and quantifies how deep (through the distance D(3), D(2) and D(1) to the respective hyper-planes) this status is. Due to the fact that the three classifiers may disagree in their respective predictions, an additional classifier to combine the individual classification into coherent predictions is necessary. This is accomplished by means of the second tier classifier (also a binary SVM based classifier), that actually provides the decision function to identify the plasma status at the time instant t.

4. APODIS TRAINING/TESTS

The APODIS predictor for ILW operation has had to be trained/tested with CFC wall discharges. All discharges between campaigns C15-C27b (years 2006-2009) have been considered as potential shots for these purposes (until the metallic wall installation, C15 was the first campaign after the last major structural modifications in JET, the replacement of the divertor). A total number of 10845 discharges have been taken into account. The first step was to remove the discharges whose signals in table 1 are either uninformative or the maximum/minimum amplitudes are outside the statistical range of the whole dataset. After this filter, the number of available discharges was 8407: 7648 non-disruptive, 521 unintentional disruptions and 238 intentional disruptions.

It is important to note that intentional disruptions may not be valid to train a predictor. The plasma evolves in a safe manner until the disruption is provoked and, depending on the type of intentional disruption, precursors may not appear.

Due to the fact that the APODIS architecture maintains good results along the time [4], two possible sets of discharges have been considered for training purposes: campaigns C15-C18 and campaigns C19-C22. The latter has produced superior results. In spite of the strong unbalanced number of disruptive and safe discharges, it was verified that balanced datasets in the training phase provide better prediction rates. As the total number of unintentional disruptions between C19 and C22 is 125, it was decided to use training datasets of 125 disruptive and 100 safe discharges (randomly selected from 2312 non-disruptive discharges). To avoid any bias in the selection of the safe pulses, 50 different training datasets have been chosen (the difference among them are the random selection of non-disruptive discharges). A multi-objective problem has been tackled in the training process: to achieve high success rate in the predictions together with a low rate of false alarms.

The first tier SVM classifiers use RBF kernels and 150 combinations of the C regularization parameter and the γ kernel parameter have been analyzed for each one of the 50 training sets (the same C and γ parameters are employed simultaneously in the 3 SVM models). Best results correspond to $C = 1000$ and $\gamma = 1$. The distances (with sign) to the respective hyper-planes are

$$D\,(M) = K\,(M) + \sum_{i=1}^{N(M)} \alpha_i(M) \cdot exp\ \{-|\mathbf{v}_i(M) - \mathbf{x}|^2\} \qquad (1)$$

where M = 1, 2, 3 is the model, K(M) is the bias of each model, N(M) is the number of support vectors in the respective SVM optimization processes, α_i(M) are the Lagrange multiplier corresponding to the support vectors, $\mathbf{v}_i$(M) are the support vectors of model M and $\mathbf{x}$ is the input feature vector. The 2nd tier decision function, F_P, uses a linear kernel with $C = 10^6$ and its mathematical formulation is

$$F_P = sign\ \{B_0 + B_1 \cdot D(1) + B_2 \cdot D(2) + B_3 \cdot D(3)$$

where D(M), M = 1, 2, 3 are the results of equation and $B_j, j = 0, \ldots, 3$ are the hyper-planes coefficients obtained in the training of the second tier classifier.

Therefore, 7500 (150x50) different predictors have been computed with thousands of training samples. To accomplish these computations in a reasonable period of time, high performance computation (HPC) has been used. Typically, 128 nodes of the CIEMAT HPC cluster (240 nodes, processors: 2 Quad-core Xeon 3.0 GHz, RAM memory 16 GB) have been used in each computation. The training of the 7500 predictors took a CPU time of 900 h for the first tier classifiers and 0.5 h for the decision function.

Shots corresponding to campaigns C23-C27b have been selected to test the predictors. All available non-disruptive discharges (3578) and all available unintentional disruptive discharges (228) have been used. The test success rate has been 93.42% (213/228), the test missed alarm rate has been 6.58% (15/228) and the test false alarm rate has been 5.11% (183/3578). The average prediction time before the disruption is 156 ms (with a maximum of 6.240s and a minimum of 1ms).

5. APODIS REAL-TIME IMPLEMENTATION IN THE RT MARTE FRAMEWORK

Although the discharges of the ILW campaigns are quite different from the ones of campaigns C19-C22 used for the training, the old models without any retraining have been applied.

The APODIS implementation in the JET RTDN was carried out under the Multithreaded Application Real-time executor (MARTe) framework [6] on a six-core x86 architecture.

APODIS has been implemented using two application threads [7]. The first one, collects the samples from the input sources. The second one receives the data from the first thread, organizes them to fit in the three-windows architecture of fig.2 and evaluates the 4 SVM classifiers.

6. APODIS RESULTS IN THE JET ILW CAMPAIGNS

As previously mentioned, the APODIS version installed to operate in the ILW campaigns has been trained with CFC wall data and no retraining has been performed. Table 2 summarizes the results. APODIS has been working even with high performance plasmas (plasma current 3.5MA, input power 25MW). Examples of these kind of discharges are Pulse No's: numbers 83479 (correctly identified as safe by APODIS) and 83480 (disruptive and predicted by APODIS 52 ms in advance).

The off-line analysis of the false alarms (Pulse No's: 82579, 82700, 83414, 83541, 83472 and 83612) determines that the first 4 discharges show thermal quenches with drops in the temperature between 400 and 1000eV that APODIS identifies as disruptions. In fact, the off-line analysis establishes that minor disruptions have taken place and the plasma has been able to recover. Concerning discharge 83472, it should be noted the presence of a faulty real-time density signal, which induces APODIS to make a mistake. Finally, the false alarm in shot 83612 is triggered during the current ramp-up around 750kA (the activation level of APODIS).

With regard to the missed alarms (Pulse No's: 82797, 82943, 83413, 83482 and 83564), they are not detected at all due to the lack of temporal resolution (32 ms is not enough). On the other hand, the estimated time to perform the prediction is about 300 μs. Due to this fact, a new APODIS version with 1 ms of resolution is ready to be implemented by means of a sliding time window mechanism. It has been verified off-line that this new version of the predictor allows detecting all the 5 disruptions missed by the slower one with 32 ms time resolution.

Figure 3 shows the cumulative percentage of the unintentional disruptions versus the warning time (the difference between the disruption time and the alarm time). A comparison is established between APODIS prediction and the JET mode lock (ML) trigger system (the mode lock trigger is based on a threshold applied on a magnetic signal). APODIS clearly outperforms the original JET predictor system. It should be mentioned that the minimum estimated time to carry out mitigation actions in JET is 30ms (green vertical line). At this time, APODIS has recognized 90% of all unintentional disruptions whereas the mode lock trigger system identifies almost 80% of them. The mean value of the APODIS prediction time is 426ms in advance to the disruption (between 1ms and 10.323s). Only 13 disruptions have been predicted with less than 30ms. The corresponding prediction time of ML is 294ms (between 1 ms and 4.850s).

7. DISCUSSION ON APODIS

APODIS is not a physics-driven system. It is an engineering system able to learn from past situations. In fact, it has demonstrated high generalization capability because it has been trained with JET CFC wall data but it has had a successful operation with the first JET ILW campaigns.

APODIS success rates and reliability during C28-C30 show that it can be used in future JET campaigns to trigger mitigation actions.

Also, it is important to mention the differences between APODIS and other machine learning disruption predictors that have been published. First, it should be noted the real-time characteristic of APODIS, that is able to follow the plasma behaviour from plasma breakdown to extinction. Second, APODIS uses a 'non-standard' set of features to represent the signals. Both temporal and frequency domains are employed. Third, the APODIS predictions take into account the plasma temporal evolution. The most recent 96 ms are considered for each prediction. And fourth, APODIS is a combination of classifiers which provides a better knowledge of the parameter space.

ACKNOWLEDGMENTS

This work, supported by the European Communities under the contract of Association between EURATOM/CIEMAT/CCFE/ENEA/IST, was carried out within the framework of the European Fusion Development Agreement. The views and opinions expressed herein do not necessarily reflect those of the European Commission.

REFERENCES

[1]. G. Matthews on behalf of JET-EFDA Contributors. 20th International Conference on PSI in Controlled Fusion Devices. Aachen, May 21st 2012.

[2]. V. Cherkassky, F. Mulier. "Learning from data". John Wiley & Sons, Inc. (1998).

[3]. G. A. Rattá, J. Vega, A. Murari, M. Johnson and JET-EFDA Contributors. Review of Scientific Instruments. **79**, 10F328 (2008).

[4]. G. A. Rattá, J. Vega, A. Murari, G. Vagliasindi, M. F. Johnson, P. C. de Vries and JET-EFDA Contributors. Nuclear Fusion. **50** (2010) 025005 (10pp).

[5]. G. A. Rattá, J. Vega, A. Murari and JET-EFDA Contributors. Fusion Engineering and Design. *http://dx.doi.org/10.1016/j.fusengdes.2012.07.002.*
[6]. A. C. Neto, F. Sartori, F. Piccolo, R. Vitelli, G. de Tommasi, L. Zabeo et al. IEEE Transactions on Nuclear Science. vol. **57**, n 2, pp. 479-486. April 2010.
[7]. J. M. López, J. Vega, D. Alves, S. Dormido-Canto, A. Murari, J. M. Ramírez et al. Proc. of the 18th REAL-TIME Conference. June 11-15, 2012. Berkeley (CA), USA.

Signal name	***Units***
Plasma current	*A*
Mode lock amplitude	*T*
Total input power	*W*
Plasma internal inductance	
Plasma density	m^{-3}
Stored diamagnetic energy time derivative	*W*
Radiated power	*W*

Table 1: List of signals to characterize the disruptive/non-disruptive status of JET plasmas.

	JET off-line classification	***APODIS predictions***
Safe	*651*	*645* (***99.08%***)
False alarms	*n/a*	*6* (***98.36%***)
Unintentional	*305*	*300* (***98.36%***))
Missed alarms	*n/a*	*5* (***1.64%***)
TOTAL	*956*	*956*
Intentional	*35*	*n/a*

Table 2: APODIS assessment: success rate, missed and false alarms during JET campaigns C28-C30.

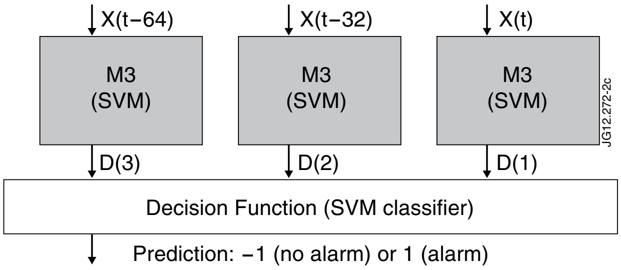

Figure 1: The training dataset is made up of circles and squares. The decision function (red dotted line) separates both classes. With SVM classifiers, the decision function in the feature space splits both classes by means of a hyper-plane (red plain line). The hyper-plane is defined by the closest samples of each class. These samples appear on the dashed lines and are called support vectors.

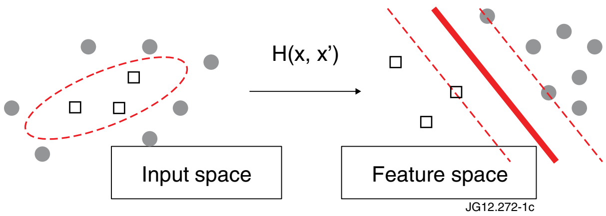

Figure 2: Multi-tier architecture of APODIS. The decision function is determined by combining the M1, M2 and M3 models. The final prediction identifies either a safe evolution (-1) or a disruptive behavior (1).

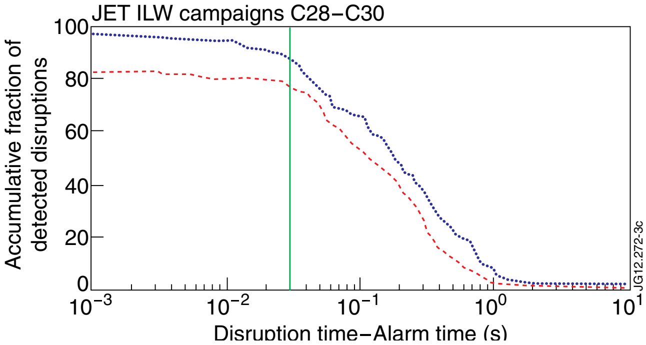

Figure 3: Comparison of the prediction times between APODIS and the JET mode lock trigger system. APODIS is faster and detects disruptions that are not recognized by ML, for example Pulse No: 82758.