

G. De Arcas, J.M. López, M. Ruiz, E.Barrera, J. Vega, G.A. Rattà,
A. Murari and JET EFDA contributors

Design of an Advanced Intelligent Instrument with Waveform Recognition Based on the ITMS Platform

“This document is intended for publication in the open literature. It is made available on the understanding that it may not be further circulated and extracts or references may not be published prior to publication of the original when applicable, or without the consent of the Publications Officer, EFDA, Culham Science Centre, Abingdon, Oxon, OX14 3DB, UK.”

“Enquiries about Copyright and reproduction should be addressed to the Publications Officer, EFDA, Culham Science Centre, Abingdon, Oxon, OX14 3DB, UK.”

Design of an Advanced Intelligent Instrument with Waveform Recognition Based on the ITMS Platform

G. De Arcas¹, J.M. López¹, M. Ruiz¹, E.Barrera¹, J. Vega², G.A. Rattà²,
A. Murari³ and JET EFDA contributors*

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

¹*Grupo de Investigación en Instrumentación y Acústica Aplicada. Universidad Politécnica de Madrid, Ctra. Valencia Km 7, 28031, Madrid, Spain*

²*Asociación EURATOM/CIEMAT para Fusión. Avda. Complutense 22, 28040, Madrid, Spain*

³*Consorzio RFX – Associazione EURATOM ENEA per la Fusione, Corso Stati Uniti 4 Padova, Italy*

** See annex of M.L. Watkins et al, “Overview of JET Results”, (Proc. 21st IAEA Fusion Energy Conference, Chengdu, China (2006)).*

Preprint of Paper to be submitted for publication in Proceedings of the
8th International FLINS Conference on Computer Intelligence in Decision and Control, Madrid, Spain.
(21st September 2008 - 24th September 2008)

ABSTRACT.

Searching for similar behavior in previous data plays a key role in fusion research, but can be quite challenging to implement from a practical point of view. This paper describes the design of an intelligent measurement instrument that uses similar waveform recognition systems (SWRS) to extract knowledge from the signals it acquires. The system is perceived as an Ethernet measurement instrument that permits to acquire several waveforms simultaneously and to identify similar behaviors by searching in previous data using distributed SWRS. The implementation is another example of the advantages that local processing capabilities can provide in data acquisition applications.

1. INTRODUCTION

Diagnostics are used in fusion experiments to gather information about the temporal evolution of the experiment and the plasma properties. Therefore for each experiment's discharge, which is identified by its pulse number, there is a set of waveforms that provide information about the temporal evolution of certain properties. All waveforms are stored in databases so they are available for offline post processing during the analysis phase. Given a single discharge, similar behaviors can be identified in previous discharges by comparing the waveforms obtained in one discharge with the ones stored in the databases. This activity plays a key role during data analysis in fusion research, but it can be quite challenging to implement from a practical point of view due to the large number of signals and discharges stored, which produces very large databases.

For example, the Joint Undertaking (JET) can produce more than 10GBytes of data during high performance experiments, each of them contributing to a database that contains already more than 40TBytes of data. Similar Waveform Recognition Systems (SWRS) have been presented as valuable tools to automate in an efficient way the identification of sets of discharges of comparable behavior [1]. Here we present the implementation of such an approach based on the Intelligent Test and Measurement System (ITMS) [2] platform for online identification. The online processing capabilities of the ITMS architecture and its versatility in terms of data acquisition, allow the implementation of a modular scalable solution. Intelligence involves the ability to learn from experience, therefore the result of this work is what is called an Advanced Intelligent Instrument (AII) [3]: a system which is capable not only of measuring, but also of identifying similar previous behaviors of the phenomena under study by comparing the data it acquires with its previous history. Section 2 provides an overview of the ITMS hardware and software architecture adapted to this specific application, while section 3 presents the SWRS model. Finally section 4 discusses the overall system functionality.

2. THE ITMS HARDWARE AND SOFTWARE ARCHITECTURE

Figure 1 shows the hardware architecture of the ITMS platform adapted to this application. The hardware consists of the following elements allocated in a PXI chassis:

1. A System Controller CPU (SCPU) which is used to configure Data Acquisition boards (DAQ),

- to acquire data from DAQ channels, to distribute the acquired data among the system's CPUs, and to process the data of the desired channels.
2. One or several DAQ boards depending on the acquisition requirements. In this case National Instruments' NI-PXI 6070E have been used, providing 16 analogue input channels with 12 bits resolution up to 1.25MHz per board, but the system architecture is open to support other boards.
 3. One or several peripheral CPU cards (PCPU) allocated in peripheral PXI slots (CC8-BLUES Pentium 3 from EKF Systems, or Inova ICP-PM-4). PCPUs are used to process the data acquired from selected channels. They increase the computing capacity of the PXI platform as compared to standard PXI systems based only in the system CPU controller [4].

The system has an external Ethernet connection so it is seen from the outside as an Ethernet instrument with a single IP address, that of the SCPU.

During measurement the SCPU retrieves the acquired data from the DAQ cards and sends it to the PCPUs for processing through the PCI bus. In this particular application the data processing results are then transferred back to the SCPU through a Local Area Network (LAN) to put them together and made them available to external systems. Depending on the system load the SCPU can also be used to process data. All the process is controlled through a set of software modules that run both in the SCPU and in the PCPUs as shown in Fig.2. Most of these modules have been developed following a client/server architecture so they are present both in the SCPU and the PCPUs.

The System Manager (SM) module controls the system behaviour by sequencing the operation of the rest of the modules and deals with the user interface. In the SCPU it can receive external commands during system setup, or data processing results from the PCPUs. It can send configuration data to the PCPUs, or pass the data results to the Results Integration (RI) module. In the PCPUs it receives the configuration from the SCPU and transmits the data processing results to the SCPU.

The system is configured using a TCP/IP based protocol. The Setup module receives this configuration and configures the Dynamic Data Processing System (DDPS) module and the data acquisition cards through the Setup&DAQ module.

The Setup&DAQ module has two different functions depending on the state of the system. During the configuration phase the module running in the SCPU configures the DAQ cards, while during the measurement phase it collects the acquired data from the DAQ cards and distributes it to the CPUs according to the system configuration. In each PCPU it receives the acquired data and passes it to the DDPS module though a set of FIFO memory buffers.

The DDPS module is in charge of running the data processing routines, in this case the SWRS as described in section 3. During the system configuration the user decides which data acquisition channels will be processed in which DDPS subsystem. For example, in this application this allows having several SWRSs in the same system for different input signals, each of them being able to run in different CPUs. The results obtained by each DDPS are sent to the Results Integration (RI) module running in the SCPU which collects and integrates them to provide a higher level of

knowledge by combining the partial results of the distributed SWRS.

The low level modules, responsible for data acquisition and distribution, have been developed as Linux kernel modules (Fedora Core 1) based on RTAI (version 3.3) using the COMEDI data acquisition driver. The rest have been developed in LabVIEW (version 8.2), except the part of the DDPS modules that deals with the continuous data collection from the FIFOs which was developed in C and integrated in LabVIEW code by using Code Interface Nodes (CINs).

3. SIMILAR WAVEFORM RECOGNITION SYSTEM DESCRIPTION

The SWRS implementation presented here follows the model described in [1]. Given an acquired signal, a SWRS is used to search for similar waveforms in previous discharges. The set of recorded signals for a given waveform is called a signal collection. The SWRS runs two processes for each new input signal: the waveform classification process and the searching process.

The waveform classification process is responsible for creating and updating the searching space, which is done following the steps shown in the left part of Fig.3. First, the signal is pre-processed by removing the offset and performing a linear interpolation to align all waveforms. This reduces the effects of different acquisition characteristics (sampling rate, signal's offset...) between discharges.

The system does not compare the waveforms itself, but feature vectors that represent them with a lower dimensionality. In this case the feature vector is obtained using the Haar wavelet transform for each waveform, which is quick to compute and keeps both time and frequency information at the same time.

In order to speed up the searching process the signal collection is organized in clusters. So the next step is to classify the feature vector in a cluster using a supervised criterion based on computing the number of features required to reach 99.5% of the feature vector Euclidean norm, which gives information on the pulse duration. The process finishes by updating the signal collection data structure, which is made of a database that stores the feature vectors and a set of files that store the pulse numbers that fall in each cluster.

The searching process searches for similar waveforms in previous discharges by comparing their feature vectors in five stages, as shown on the right side of Fig.3: (1) the waveform is preprocessed to obtain its feature vector, (2) the later is classified into a cluster, (3) all the pulse numbers in that cluster and the two adjacent ones are retrieved, (4) similarity is computed between all feature vectors retrieved, and (5) the results are used to sort the pulse numbers in descending order. The similarity, S_{uv} , between two feature vectors u_w and v_w is measured using the absolute value of their normalized inner product avoiding dependency on gain factors and/or polarity,

$$S_{uv} = \frac{|u_w \cdot v_w|}{\|u_w\| \cdot \|v_w\|}, \quad 0 \leq S_{uv} \leq 1 \quad (1)$$

It must be noted that this technique retrieves the most similar waveforms to a given one, but does not imply that the signals are almost equal ($S_{uv} \approx 1$).

4. SYSTEM FUNCTIONALITY

A software application allows the user setting up the system for a new acquisition, and/or viewing the results after each discharge. The setup options are divided in two categories: acquisition setup allows selecting all acquisition parameters, such as number of channels, polarity, gain, sample rate, etc; SWRS setup allows creating or removing signal collections and choosing which SWRS are used for each acquisition channel. A new signal collection can be created empty if there is no previous information on that waveform, or it can be initialized by uploading the signals collection's data structure. In both cases, the system's waveform classification process, described in section 3, guarantees its continuous learning as new signals are added to the signal collection for each pulse discharge. This scheme supports diagnostic's evolution and at the same time takes advantage of previous knowledge.

Once the user has defined which signals are going to be acquired and has assigned them to a SWRS module, the system is ready to use. After each acquisition each SWRS provides local information of the most similar previous discharges according to each signal. By combining the results of all SWRS together the system provides more precise information.

ACKNOWLEDGMENTS

This work has been partially supported by the Spanish Ministry of Science and Education under the research projects N° DPI200606624 and N° ENE2004-07335 and also by grant JC2007-00084 of the José Castillejo Program.

REFERENCES

- [1]. J. Vega et al. Data mining technique for fast retrieval of similar waveforms in Fusion massive databases, *Fusion Eng. Design* (2007), doi:10.1016/j.fusengdes.2007.09.011
- [2]. E. Barrera et al. PXI-based Architecture for Real Time Data Acquisition and Distributed Dynamic Data Processing. *IEEE TNS* 53-3 (2006) 923-926
- [3]. M. Ruiz et al. Data reduction in the ITMS system through a data acquisition model with self-adaptive sampling rate, *Fusion Eng. Design* (2007), doi:10.1016/j.fusengdes.2007.09.015
- [4]. M. Ruiz et al. Distributed real time data processing architecture for the TJII data acquisition system. *Review of Scientific Instruments* **75**-10 (2004) 4261-4264.

