LIT Dataset data acquisition specification

Version 3 (2019-01-16)

1. Introduction

This document aims at specifying the formats for the data that will be acquired in order to build the LIT data set. These data formats concern the way the acquired data is stored and also the auxiliary information that will be useful on building the data set for power signature identification.

The data will be collected and stored by an Acquisition System (AS). The following requirements are defined for the AS:

- R1: The AS shall collect voltage and current samples from the electrical circuit at which a set of suitable power loads for the data base construction will be switched on or off;
- R2: The AS shall collect samples at 15360 +- 0,5% Hz;
- R3: The AS shall obtain absolute timestamps;
- R4: The AS shall allow identifying the absolute time at which every set of voltage and current samples has been acquired;
- R5: The AS shall detect a set of load switching events;
- R6: The AS shall annotate the precise sample set during which the occurrence of a load switching event has been detected;
- R7: The AS shall annotate the identification of the load that generated a specific load switching event.

The task of acquiring a set of voltage and current samples, while generating controlled events on a set of loads attached to the circuit(s) being monitored is called an *acquisition session*. This concept is used to help on organizing the acquired data for construction of the LIT dataset.

The next sections describe how the voltage and current data, as well as the load events, are acquired, processed and stored during an acquisition session.

2. File formats description

In accordance with the previously defined requirements, the file formats for the AS collected data have been designed. This section describes the file formats in detail.

Every acquisition session performed by the AS creates and generates the set of files described below in a file system. *NNN* corresponds to a session number, which is automatically generated by the AS and incremented, starting from 000, for each new session.

- Samples_*NNN*.tdms: file in TDMS format [1] containing the acquired samples for session *NNN*. This file is generated only when the AS is a NI MyRIO device.
- Samples_*NNN*.bin: file in raw binary format containing the acquired samples for session *NNN*. This file is generated only when the AS is not a NI MyRIO device.
- Samples_*NNN*.events: file containing a timestamp and the annotated list of events for the corresponding session, as described in Section 2.2.
- Samples_*NNN*.config: file containing configuration data for the corresponding session, as described in Section 2.3.

The following subsections detail the sample, event and config file formats, respectively.

2.1 Sample file formats (Samples_*NNN*.tdms or Samples_*NNN*.bin)

The sample data are stored the same way for both the TDMS and binary formats. Therefore, this section will focus on the sample data organization, without addressing specific details of the file formats, e. g., the TDMS headers format.

On every sampling interval, the AS collects a sample set. This sample set is composed of a number Nv of voltage sensor and a number Nc of current sensor values. The values of Nv and Nc correspond to the *NumOfVSensors* and *NumOfISensors* parameters described in Section 2.3.

Each sample of a set is stored as a Nb-bit unsigned integer number, where Nb corresponds to the *SampleWidth* parameter described in Section 2.3. Figure 1 and Table 1 detail the bitmap for a sample set.

Num bits	DW-Nb-1	Nb	1	DW-Nb-1	Nb	1	...	DW-Nb-1	Nb	1
Description	Stuffing	V_0 Sample	PPS flag	Stuffing	V_1 Sample	PPS flag	....	Stuffing	V_{Nv-1} Sample	PPS flag

DW-Nb-1	Nb	1	DW-Nb-1	Nb	1	...	DW-Nb-1	Nb	1
Stuffing	I_0 Sample	PPS flag	Stuffing	I_1 Sample	PPS flag	....	Stuffing	I_{Nc-1} Sample	PPS flag

Figure 1 – Format of a sample set

Table 1 – Sample set field descriptions

Field	Description

V$_i$ sample	Voltage signal sample for the *i-th* voltage sensor, as a *Nb*-bits unsigned integer. This value is obtained by converting the sample generated by the A/D converter of the AS.
I$_i$ sample	Current signal sample for the *i-th* voltage sensor, as a *Nb*-bits unsigned integer. This value is obtained by converting the sample generated by the A/D converter of the AS after the current provided by a Current Transformer has been converted to a voltage signal.
PPS flag	1 when the PPS signal generated by the GPS is detected during the acquisition of this specific sample set, 0 otherwise (see Section 2.2 for further information about the use of the GPS device).
DW	Data width for every sample value. DW is the minimum multiple of 16 which is greater than the value of *Nb* plus 1. This allows the file data to be aligned to multiples of half-word boundaries.
Stuffing	Stuffing bits set to 0.

As a consequence of the PPS flag and the stuffing bits, the width of a sample set is equal to DW times the total number of voltage and current sensors.

2.2 Event file format (Samples_*NNN*.events)

The event file is split into two sections:

- Section 1: timestamp, in text format, corresponding to the number of seconds elapsed since 12:00 AM of 01/01/1970 at the instant when the first valid sample (i. e. correctly synchronized with the absolute time) is acquired.

The timestamp is generated by AS by decoding a specific time message generated by a GPS module. This module also provides a digital pulse-per-second (PPS) signal, which allows the AS synchronizing the transitions of seconds, in absolute time, with a typical jitter of 100 ns.

- Section 2: events registry.

Every event related to a load switching is stored as a record into the events registry. The record corresponds to a line of text with the following format:

YYYY:MM:DD:HH:MM:SS, sample_index, event_ID

Table 2 describes the record field.

Table 2 – Record field descriptions

Field	Description
YYYY:MM:DD:HH:MM:SS	Absolute time with resolution of seconds, during which the event has occurred. *YYYY:MM:DD* corresponds to the date (4 digits for year, 2 digits for month and 2 digits for day) and *HH:MM:SS* corresponds to time (2 digits for hour, 2 digits for minute and 2 digits for second).
sample_index	Index of the sample at which the event has occurred, starting on the last PPS annotated sample pair (PPS flag set to 1). As the sampling frequency is 15384 Hz, valid values for *sample_index*

	range from 0 to 15383.
event_ID	Event identifier, as defined by an event identifier table (obs: on/off events for the same load type are identified by different event IDs).

2.3 Config file format (Samples_*NNN*.config)

The config file is organized as a textual list of parameters, one per line. Each parameter is given a configuration value in the following format:

Parameter=value

Table 3 shows the configuration parameters and a description of the valid values.

Table 3 – Configuration parameters and description

Parameter	Value description
SampleWidth	Number of valid bits for every sample
NumOfVSensors	Number of sampled voltage sensors
NumOfISensors	Number of sampled current sensors
Kv	List of comma-separated floating-point values for the scale factors of each of the *NumOfVSensors* voltage sensors
Ki	List of comma-separated floating-point values for the scale factors of each of the *NumOfISensors* current sensors
ZeroOffsetV	List of comma-separated values for the digital (decimal) values corresponding to the 0 V input for each of the *NumOfVSensors* voltage sensors
ZeroOffsetI	List of comma-separated values for the digital (decimal) values corresponding to the 0 A input for each of the *NumOfISensors* current sensors
Format	0 = binary, 1 = TDMS
GridFrequency	Frequency of the power grid being sampled in Hz
SamplesFrequency	Number of samples per second
SessionDescription	Session description as free text, enclosed with double quotes

Example:

For the acquisition session with the following setup:

2 voltage channels, 2 current channels, 12-bits unsigned integer per sample, voltages ranging from -10 to +10 V, currents ranging from 0 to +8 A, samples stored in a TDMS file, 60 Hz, 15384 samples per second.

The config file would have the following contents:

SampleWidth=12

NumOfVSensors=2

NumOfISensors=2

Kv= 0.0048828125,0.0048828125

Ki=0.001953125,0.001953125

ZeroOffsetV=2048,2048

ZeroOffsetI=0,0

Format=1

GridFrequency=60

SamplesFrequency=15384

SessionDescription="

This is an example of configuration file for an acquisition session. Kv and Ki values allow converting the digital V and I values to analog V in the range (-10,+10) and to analog I in the range (0, +8)"

3. Event detection and identification

In order to fulfill requirements R5 and R7, the AS shall detect and identify the occurrence of load events. This is performed by means of notifications sent by an external system (e. g. the hardware responsible for controlling or monitoring the loads that generate the events) and captured by a set of AS inputs.

Event detection and identification in the AS is implemented as follows:

- 8 digital inputs are used to receive the event identification (EID) value;
- 1 digital input is used to receive the event detection notification (EDN);
- The EID is latched on the rising edge of the EDN.

The 8-bit length EID allows defining 256 different event numbers. Each event number identifies an event (e. g. ON or OFF) and the corresponding load. For example, the ON and OFF events of a PC monitor could be identified by EIDs 130 and 131, respectively.

4. Voltage and current value calculation

The actual value of a voltage sample stored in the sample file is calculated as follows:

$$V_a = (V_i - ZeroOffsetV_i) * Kv_i$$

Table 4 – Parameters for voltage calculation

Parameter	Description
V_a	Actual voltage value
V_i	Digital (decimal) value of the voltage signal sample for the *i-th* voltage

	sensor
ZeroOffsetV$_i$	Digital (decimal) value corresponding to the 0 V input for the *i-th* voltage sensor (see Section 2.3)
Kv$_i$	Scale factor for the *i-th* voltage sensor (see Section 2.3)

Similarly, the actual value of a current sample stored in the sample file is calculated as follows:

$$I_a = (I_i - ZeroOffsetI_i) * Ki_i$$

Table 5 – Parameters for current calculation

Parameter	Description
I$_a$	Actual current value
I$_i$	Digital (decimal) value of the current signal sample for the *i-th* current sensor
ZeroOffsetI$_i$	Digital (decimal) value corresponding to the 0 A input for the *i-th* current sensor (see Section 2.3)
Ki$_i$	Scale factor for the *i-th* current sensor (see Section 2.3)

For the session configuration example shown in Section 2.3, a voltage sample with digital value equal to 1035 would correspond to:

$$V_a = (1035 - 2048) * 0.0048828125 = \textbf{-4.9462890625 V}$$

5. References

[1] National Instruments, "The NI TDMS File Format". Available in http://www.ni.com/white-paper/3727/en/