IEEE PSRC WG H26

COMTRADE 2013 Conformity Test Plan

Report to the Main Committee
May 2015

Assignment:

Develop a plan that can be used to test COMTRADE files for conformity to the IEC 60255-24 Ed 2.0 2013-04 / IEEE Std C37.111™ standard.

Contact Information:

Any questions or comments on the testing procedure described in this report may be sent to COMTRADE2013@gmail.com.

Working Group Members:

J. Rick Cornelison (Chair) **Jay Gosalia (Vice-chair)**

IEEE PSRC WG Members:

Eric Allen, Ratan Das, Jim Hackett, Yi Hu, Mladen Kezunovic, Theo Laughner, Amir Makki, Jose Mar, Greg Rzepka, Mark Taylor, Benton Vandiver, Jun Verzosa

IEEE Liaisons:

Lloyd Green, Ravi Subramaniam

Contents

1 Introduction

The common themes of the COMTRADE standard evolution over the last 20 years are innovation, cooperation, and success. The COMTRADE standard evolved in four stages: the first stage is represented by the IEEE 1991 standard, the second stage is represented by the IEEE 1999 revision, the third stage is represented by the consequences of the 2003 Northeast blackout, and the fourth stage is represented by the 2013 IEEE/IEC Dual Logo COMTRADE standard. The goal of the first COMTRADE standard [1] was to assist relay engineers in evaluating relay performance based on Digital Fault Recorder (DFR) data. The first standard was quite successful: it was accepted by many relay manufacturers and utility engineers who recognized the value in sharing and using data from many devices scattered over a large geographical area, including simulation laboratories, test laboratories, and field installations. Noting the success of the COMTRADE standard in North America, users in other countries also began using the standard, resulting in a COMTRADE version of the standard coming from European manufacturers with 1993 identified as the standard year. Users were attracted to the simplicity and the interoperability provided by the standard.

During the 1990s, many digital relays capable of recording transient data were installed for the first time. Subsequently, the field of transient data recording was now open to providers of digital relays as well and hence a considerable number of new operating nuances were introduced. The Power System Relaying Committee kept close track of this evolution and delivered a revised IEEE COMTRADE standard in 1999 [2]. The main objective of the revised standard was to reinforce the basic concepts of transient data representation and to accommodate a growing set of innovative discoveries such as “dynamic sampling” and “information files”.

International acceptance of COMTRADE was achieved throughout the community of users and providers (including DFR, digital relay, and third party providers) around the turn of the millennium. This acceptance led to another considerable number of new providers joining the community, including providers of digital meters, phasor measurement units, and remote terminal units. This in turn added another considerable number of derivatives, deviations, and interpretations of the standard. For example, some providers began using COMTRADE to exchange non-transient information. IEC adopted the revised COMTRADE standard in 2001 [3].

On August 14, 2003, a large part of the northeastern United States and part of Canada experienced a blackout that affected over 50 million people and required several days for full restoration of the system. The North American Electric Reliability Council began an investigation into the root causes behind the Northeast blackout, with the aim of making recommendations to help avoid such events in the future. Thousands of transient data files were collected from many of the utilities involved. The collected files revealed several fundamental problems that seriously affected the investigative process:

- The collected files were in multiple formats, many of which were proprietary. The investigative team had to use multiple types of display and analysis programs, which slowed the analysis process and hindered their ability to synchronize data files and perform end-to-end studies.
- The collected files did not share a common naming convention: which made it difficult to discern which files came from particular utilities and also which files were captured by particular devices. The lack of a common naming practice became a serious hindrance in the investigation.

- The collected files lacked a formatted field for specifying whether the time tags were based on local time or UTC (UTC stands for Coordinated Universal Time, and UTC without offset is called “Zulu” time).
- The collected files lacked a common set of formatted fields for specifying whether or not the time tags were properly synchronized.
- The lack of a combined file format resulted in added complications of managing and keeping track of up to 4 separate files for each COMTRADE record.

Since 2002, two Power System Relaying Committee working groups (H5D and H4) have been actively collecting lessons learned, developing solutions, and carefully revising the original text of the 1991 and 1999 standards, an effort which has resulted in the 2013 IEEE/IEC Dual Logo COMTRADE Standard [4]. The standard contains a number of useful additions and revisions including:

- a number of new fields and data types designed to support the standard’s growing use and expanding scope;
- a new single file structure designed to simplify the management and tracking of large quantities of COMTRADE records;
- revised text that removes a number of obsolete restrictions, such as the restriction of filenames to 8 characters (Current naming practices, such as the COMNAME format, allow for filenames of up to 253 characters in length, but the use of no more than 64 characters is recommended to ensure compatibility across various types of operating systems.); and
- some of the non-critical fields in the configuration section have been made critical for better understanding and use of the data in the COMTRADE file.

With the changes in the 2013 COMTRADE standard, it was felt necessary to develop a test plan to ensure conformity of COMTRADE files with the standard.

2 Purpose

The purpose of this test plan is to ensure a common testing methodology for COMTRADE files, irrespective of the sources. This will ensure compatibility of files from different sources and compliance to the 2013 COMTRADE standard.

3 Source of Files

Users define the need for transient data and use of the COMTRADE standard. The original source for such data and the main reason COMTRADE exists are Digital Fault Recorders (DFR). The primary purpose of such devices is the recording of transient waveforms of power system events; however this has expanded since that first version of COMTRADE in 1991. DFR type data are now a common capability of nearly all intelligent electronic devices that are applied to a protection and control scheme, including early digital relays and meters where the event files can be downloaded and converted to COMTRADE files.

Today the physical sources of COMTRADE data come from any industry device that has analog and digital time-based data that need to be recorded and transportable in a standard file format. The COMTRADE 1999 standard transcended the original power system use case to be adopted by other technicians, engineers, and scientists for all types of testing and investigations where time-based sampled data were needed. A few examples are: system using analog gauges, chart recorders, numerical sensors (temperature, pressure, strain, etc.) and encoders, all types of meters, binary status, and voltage or current analog/digital samples. Each modern protection relay vendor has a software tool that can manipulate event data or power quality data in the vendor's native format and provide these data as a COMTRADE file.

In addition to the physical sources, there are now software-based simulation sources of COMTRADE data. These programs generally have a primary purpose of simulating a physical system: EMTP – Electro Magnetic Transient Program and ATP – Alternative Transient Program are used for power system transient simulations. But there are other general purpose programs like MathCad or MatLab which are used for algorithm development and engineering calculations that can generate sampled data which are often visualized and exported as COMTRADE files.

There are numerous COMTRADE viewers which can import the COMTRADE data files to visualize them, but many can also convert other sampled data formats into COMTRADE data files or modify COMTRADE files and re-export them tailored to a user's new requirement. Many such data formats now involve streaming data like PMU – Phasor Measurement Unit streams or IEC 61850-9-2 Sampled Values streams for which data mappings have been defined to put them in the COMTRADE file format. This is primarily to leverage the established base of viewers and analysis software for COMTRADE files.

The need for more dynamic and system testing of industry devices in all protection automation and control schemes has expanded the use of COMTRADE to provide the required test cases. All types of testing – acceptance, commissioning, routine, and especially trouble shooting of mis-operations have utilized COMTRADE as a source for those dynamic system events. Most test set manufacturers directly support the use and conversion of COMTRADE files from all the above sources in order to provide these dynamic testing abilities.

This extensive usage of the standard drives its continued revision and resulted in the COMTRADE 2013 changes to support all of the evolving applications that COMTRADE now supports worldwide.

4 Common Errors

This section discusses common errors in the CFG and DAT files. The HDR and INF files are not discussed because they are optional, open format files. The common errors are categorized as formatting, interpretation, and/or designation errors.

4.1 Formatting Errors

These are mainly format adherence errors that usually render the files unusable or unreadable by display or automated analysis applications. The typical symptoms of formatting errors are when the applications abort, crash or get caught in endless loops. The symptoms appear upon attempting to open the files. Examples of common formatting errors include:

- Missing or extra data separators (total comma count mismatch).
- Missing or redundant <CR/LF> at the end of each line of text.
- Missing <EOF> marker at the end of the ASCII type data file.
- Data separator or <CR/LF> used in Binary type data file.
- The defined number of analog or digital channels identified in line 2 of the CFG file does not match the number of channels subsequently listed.

4.2 Interpretation Errors

These are mainly representation errors that may lead to misinterpretation of data. These errors are critical issues because engineers may take the wrong actions causing additional problems. Such actions may include sending restoration crews to the wrong fault location or providing relay technicians with the wrong settings. These errors are also elusive because they do not actually prevent numerical data from being read. Applications may open such files and display them without any notice of anything wrong. The typical symptoms of these errors become apparent when the values are checked. For example, the pre-fault line to neutral voltage for a 345 kV bus should be approximately 200 kV, but the values displayed or calculated by the applications show otherwise. Common examples of conversion errors include:

- Analog channel units not defined properly or missing.
- Primary types defined but data provided in secondary values, or vice versa.
- Primary and secondary ratios not defined in post 1991 revisions.
- Scale factor not adjusted for RMS calibrated data.
- Max and min ranges do not match raw data ranges (accuracy undeterminable).
- Line frequency set to 50Hz while raw data is 60Hz, or vice versa.
- Date and time format not compliant with the revision year (month and day reversed).
- Trigger date and time occurs earlier than start date and time.
- Normal state of digital channels not defined.
- Real data types used for pre-2013 revisions.

4.3 Designation Errors

These are mainly labeling errors that may lead to further misinterpretation of the data. These errors are also elusive because they also do not actually prevent numerical data from being read. Typical symptoms of these errors become apparent when the labels are checked. For example, the Phase A over current element asserts when Phase B has an over current condition. Such errors may lead the analyst to set the wrong elements. Designation errors may also induce rotational issues causing reversal of negative and positive sequence components. Common examples of designation errors include:

- Digital channels listed in reverse order in Binary type data file.
- Trailing commas used for missing analog data instead of 99999 for pre-2013 revisions.
- Raw data range exceeds 99999 for pre-2013 revisions.
- Revision year is not defined and the format is not the 1991 revision.
- Revision year listed as 1997 or other year for which no COMTRADE revision exists.

5 Testing of Four Separate Files

Unless stated otherwise, the outcome of all of the testing described below should be either pass or fail. Any single failure should result in the file failing to pass the test. If there is any ambiguity, a comment should be generated.

5.1 CFG File

The configuration file is an ASCII or UTF-8 text file that provides the information necessary for a human or a computer program to read and interpret the data values in the associated data file. The configuration file has a predefined, standardized format. The file is divided into lines and the format and exact order of each line are specified. Testing to be done is as follows:

1. Check the filename extension, it should be **.CFG** extension.
2. For each line, check for proper line ending and that each field in a line is separated from the preceding field by a comma (“,”).
3. Check for each of the three fields in line 1
    a. Field 1 (station name) should be 1-64 alphanumeric characters (Standard says, Minimum character = 0 which needs to be fixed during the next revision).
    b. Field 2 (recording device identification) should be 1-64 alphanumeric characters (Standard says, Minimum character = 0 which needs to be fixed during the next revision).
    c. Field 3 (standard revision year) should read as 2013.
4. Check for each of the three fields in line 2
    a. Field 1 (total number of channels) should be 1-6 integer characters and must equal to the sum of numbers in field 2 and field 3.
    b. Field 2 (number of analog channels) should be 2-7 alphanumeric characters with the last character being “A” while other characters before it must be numeric integer characters.
    c. Field 3 (number of digital channels) should be 2-7 alphanumeric integer characters with the last character being “D” while other characters before it must be numeric characters.
5. Check each field of the subsequent lines containing analog data information (one line per analog channel) as follows:
    a. Field 1 (analog channel index number) matches with the analog channel number and 1-6 integer characters.
    b. Field 2 (channel identifier) should be 1-128 alphanumeric characters.
    c. Field 3 (channel phase identification) should be 0-2 alphanumeric characters which mean that this field is optional.
    d. Field 4 (circuit component being monitored) should be 0-64 alphanumeric characters which mean that this field is optional.
    e. Field 5 (channel units) should be 1-32 alphabetic characters. Please refer to [5], [6], and [7] for allowed units and letter symbols for physical quantities. The word “NONE” is to be used for unit-less values.
    f. Field 6 (channel multiplier [“a”]) should be 1-6 characters, real and numeric; standard floating point notations are acceptable.
    g. Field 7 (channel offset [“b”]) should be 1-6 characters, real and numeric; standard floating point notations are acceptable.
    h. Field 8 (channel time skew) should be 1-32 characters, real and numeric; standard floating point notations are acceptable.
    i. Field 9 (range minimum for data values) should be 1-13 characters, real or integer and numeric; standard floating point notations are acceptable; 3.4028235E38 ≤ Field 9 ≥ -3.4028235E38.
    j. Field 10 (range maximum for data values) should be 1-13 characters, real or integer and numeric; standard floating point notations are acceptable; 3.4028235E38 ≤ Field 10 ≥ -3.4028235E38; Field 10 ≥ Field 9.
    k. Field 11 (VT/CT primary factor) should be 1-32 characters, real and numeric.
    l. Field 12 (VT/CT secondary factor) should be 1-32 characters, real and numeric.
    m. Field 13 (primary/secondary data identifier) should have one of the characters: “P”, “p”, “S” or “s”
6. Check each field of the subsequent lines containing digital data information (one line per digital channel) as follows:

   a. Field 1 (digital channel index number) matches with the digital channel number and 1-6 integer characters.
   b. Field 2 (channel identifier) should be 1-128 alphanumeric characters.
   c. Field 3 (channel phase identification) should be 0-2 alphanumeric characters which means that this field is optional.
   d. Field 4 (circuit component being monitored) should be 0-64 alphanumeric characters which means that this field is optional.
   e. Field 5 (normal state) should be integer and numeric and should be either “0” or “1”.
7. The line frequency is the next line after the digital data information. The single field in this line should be 1-32 characters, real and numeric, standard floating point notations are acceptable. (Standard says, Minimum character = 0 which needs to be fixed during the next revision).
8. The next section (sampling rates) should be checked as follows:
   a. The first line of this section (number of sampling rates) should be 1-3 characters, integer and numeric, $999 \geq \text{Line } 1 \geq 0$
   b. This line should be followed by one line for each sampling rate, or a single line if the number of sampling rates is zero. Each of these lines should have two fields:
      i. Field 1 (sample rate) in each line should be 1-32 characters, real and numeric, standard floating point notation may be used. If the number of sampling rates is zero, this field should be zero.
      ii. Field 2 (end sample) in each line should be 1-10 characters, integer and numeric, $9999999999 \geq \text{Field } 2 \geq 1$
9. Check the next two date and time lines as follows:
   a. Field 1 (time of first data value) has three subfields which are separated by a “/”
      Subfield 1: 1-2 integer characters, $31 \geq \text{Subfield } 1 \geq 01$.
      Subfield 2: 1-2 integer characters, $12 \geq \text{Subfield } 2 \geq 01$.
      Subfield 3: 4 integer characters, $9999 \geq \text{Subfield } 3 \geq 1900$.
   b. Field 2 (trigger time) has three subfields which are separated by a “:”
      Subfield 1: 2 integer characters, $23 \geq \text{Subfield } 1 \geq 00$.
      Subfield 2: 2 integer characters, $59 \geq \text{Subfield } 2 \geq 00$.
      Subfield 3: has two subfields separated by a “.”
         Sub-subfield 1: two integer characters,
         Sub-subfield 2, 6-9 integer characters,
         $59.999999999 \geq \text{Subfield } 3 \geq 00.000000$.
10. The next line (data file type) will have 5-8 alphanumeric characters with one of the following: “ASCII”, “binary”, “binary32” or “float32” (Standard says, alphabetic characters which needs to be fixed during the next revision).
11. The next line (time stamp multiplication factor) should have 1-32 real, numeric characters, standard floating point notations can be used.
12. The next line (time codes) should have two fields, each with 1-6 alphanumeric characters.
13. The next line should have two fields as follows:
   a. Field 1 (time quality code) should have one Hex character 0-F.
   b. Field 2 (leap second indicator) should have one integer numeric character, $3 \leq \text{Field } 2 \geq 0$

5.2 INF File

The .INF file provides for the exchange of information regarding the event recorded in the COMTRADE record that may enable enhanced manipulation or analysis of the data. The information file is an ASCII or UTF-8 text file that is in a computer-readable specified format. The file contains information readable by the general user and information specific to a given class of users, which may be unreadable to the general user. These two types of information are classified as public and private, respectively, and reside in separate sections of the file. Data stored in the information file shall be stored in a public section whenever a suitable section is defined. If a suitable predefined public section is not available, a private section may be used. The entries shall conform exactly to the format defined below so that the data can be read by a computer program. Public sections contain information in a form that can be used by equipment and/or software made by more than one manufacturer. Private sections contain manufacturer-specific information that is only useful with a specific vendor's software or hardware, or that is in a format unique to that manufacturer. Multiple private sections are allowed per manufacturer, and a single information file may contain private sections from several manufacturers. It is anticipated that manufacturers will generate private sections for specific purposes. If two or more manufacturers use similar private sections, a common form of the private section could be approved for use as public sections in future revisions of this standard. Structure, syntax and more details about the information file/section from the standard is provided in Appendix B for ready reference. Testing to be done as follows:

1. Check for the file extension; it should be **.INF**.
2. The file only contains ASCII or UTF-8 supported characters.
3. No leading spaces are in any line.
4. File does not contain any user-added end-of-file marker such as "1A" HEX
5. File size is less than or equal to 64k.
6. Checking for Public and Private Section headers:
    a. Public and private section names are delimited by square brackets.
    b. Each section name is contained in one line.
    c. Each line is terminated with a <CR/LF>.
    d. Each section name's first character is a letter character, not a number or a symbol.
    e. Section name must start with the word "Public" or, for private sections, a word clearly representing the organization to which the section belongs, followed by exactly one space, then followed by any number of words identifying the section.
    f. Individual words in proprietary company or organization names or trademarks comprising more than one word shall be concatenated by deleting the space between the words, or, to improve readability, by substituting the underline space character "_" for the space.
    g. Section headings after the first section heading shall be separated from the preceding section header or entry lines by an empty line.
7. Checking for Public and Private Section entry lines:
    a. An entry line must start with one word 3–32 characters long followed by an equal (=) sign. The first word is the "Entry Name."
    b. The entry name shall be meaningful when read in conjunction with the section name.
    c. The entry name can contain any printable characters with ASCII values between 33 and 127 decimal and corresponding UTF-8 characters.
8. Checking for Value String in Entry lines:
    a. Multiple data items are separated by commas.
    b. Numeric values begin immediately after the equal sign or comma delimiter with no leading space.
    c. Text strings that include a space after the equal sign or comma delimiter shall include the space as part of the value.
9. Checking for comment section and comment lines:
    a. An entry line starting with a semicolon is called comment line and should be skipped by file reading algorithms.

   b. When a section heading is commented out, all entry lines within the section should also be commented out.
10. Checking for the Public Record Information Section:
   a. Public Record Information Section should have the first 8 lines as listed below:
      i. [Public Record_Information]<CR/LF>
      ii. Source=Value<CR/LF>
      iii. Record_Information= Value <CR/LF>
      iv. Location= Value <CR/LF>
      v. max_current=Value<CR/LF>
      vi. min_current=Value <CR/LF>
      vii. max_voltage= Value <CR/LF>
      viii. min_voltage=Value <CR/LF>
      ix. EventNoteCount=Value <CR/LF>
   b. Line 1 is the header line and should contain a text string of "Public Record_Information".
   c. Value in line 2 should be an alphanumeric string with multiple data items separated by commas.
   d. Value in line 3 should be an alphanumeric string with multiple data items separated by commas. The first 4 data items are publicly defined as follows:
      Value1: Fault, Unknown, Misoperation, Close, Trip, Reclose, Power Swing, Simulation
      Value2: AG, BG, CG, ABCG, AB, BC, CA, ABC, or any similar series of phase identifier such as 12N, RS, etc.
      Value3: Any other text string not being a variation of one of the above that helps describe the event.
      Value4: Any other text string being an identifier for a unique device or type of device.
   e. Value in line 4 should be an alphanumeric string with multiple data items separated by commas. The first 2 data items are publicly defined as:
      Value1: Should be a real number
      Value2: Should be Miles, kilometers, percent of line, percent of setting, or Ohms.
   f. Value in lines 5, 6, 7 and 8 should be a real number.
   g. Line 9 is required only if the Event Information Sections are included. Value should be an integer number and equal to the total public event information in INF file.
11. Check for the Public Event Information Section:
   a. Public Event Information Section should have 8 lines listed below:
      i. [Public Event_Information_#n] <CR/LF>
      ii. Channel_number=Value<CR/LF>
      iii. max_value=Value <CR/LF>
      iv. min_value=Value <CR/LF>
      v. max_sample_number=Value <CR/LF>
      vi. min_sample_number=Value <CR/LF>
      vii. Sample_number_Text_#n=Value1, Value2 <CR/LF>
      viii. Sample_number_Text_#n=Value1, Value2 <CR/LF>
   b. Line 1 is the header line and should be the string of "Public Event_Information_#n", where n should be a positive integer and equal to the value of EventNoteCount in the line9 of Public Record Information Section.
   c. Value in line2 should be a string of a positive integer number.
   d. Value in lines 3 and 4 should be a string of a real number.
   e. Value in lines 5 and 6 should be a string of a positive integer number.

f. n in lines 7 and 8 should be 1-2 integer characters with $1 \leq n \leq 99$, Value1 should be a string of positive integer number and Value2 should be an alphanumeric string.

12. Check for the Public File Description Section:
    a. Only one Public File_Description section is allowed per record.
    b. Public File Description Section should conform to the following format:

        [Public File_Description]<CR/LF>
        Station_Name=Value<CR/LF>
        Recording_Device_ID=Value<CR/LF>
        Revision_Year=Value<CR/LF>
        Total_Channel_Count=Value <CR/LF>
        Analog_Channel_Count=Value<CR/LF>
        Status_Channel_Count=Value<CR/LF>
        Line_Frequency=Value<CR/LF>
        Sample_Rate_Count=Value<CR/LF>
        Sample_Rate_#1=Value<CR/LF>
        End_Sample_Rate_#1=Value<CR/LF>
        Sample_Rate_#n=Value<CR/LF>
        End_Sample_Rate_#n=Value<CR/LF>
        File_Start_Time=Value<CR/LF>
        Trigger_Time=Value<CR/LF>
        File_Type=Value<CR/LF>
        Time_Multiplier=Value<CR/LF>

    c. Line 1 is the header line and n is an integer number with $1 \leq n \leq$ the number of sampling rates for the record.
    d. The entries for “Value” should follow the rules for the equivalent data as specified in section 5.1 for the .CFG file.
13. Check for the Public Analog Channel Section:
    a. One public channel description section is required for each analog channel of the record.
    b. Public Analog Channel Section should have 13 lines defined as below:

        [Public Analog_Channel_#n]<CR/LF>
        Channel_ID=Value<CR/LF>
        Phase_ID=Value<CR/LF>
        Monitored_Component=Value<CR/LF>
        Channel_Units=Value<CR/LF>
        Channel_Multiplier=Value<CR/LF>
        Channel_Offset=Value<CR/LF>
        Channel_Skew=Value<CR/LF>
        Range_Minimum_Limit_Value=Value<CR/LF>
        Range_Maximum_Limit_Value=Value<CR/LF>
        Channel_Ratio_Primary=Value<CR/LF>
        Channel_Ratio_Secondary=Value<CR/LF>
        Data_Primary_Secondary=Value<CR/LF>

    c. Line 1 is the header line and n is an integer number with $1 \leq n \leq$ the count of analog channels for the record.
    d. The entries for “Value” shall follow the rules for the equivalent variables as specified in section 5.1 for the .CFG file.
14. Check for the Public Status Channel Section:
    a. One public channel description section is required for each status channel of the record.
    b. Public Status Channel Section should have 5 lines defined as below:

[Public Status_Channel_#n]<CR/LF>
Channel_ID=Value<CR/LF>
Phase_ID=Value<CR/LF>
Monitored_Component=Value<CR/LF>
Normal_State=Value<CR/LF>

c. Line 1 is the header line and n is an integer number with $1 \leq n \leq$ the count of status channels for the record.
d. The entries for "Value" shall follow the rules for the equivalent variables as specified in section 5.1 for the .CFG file.

5.3 HDR File

The header file is an optional ASCII or UTF-8 text file for the storage of supplementary narrative information. The file is intended to be printed and read by reader, not manipulated by an application program. It could include any chosen information and in any order desired. Testing to be done as follows:

1. Check for the file extension; it should be **.HDR**.
2. The header file is a freeform ASCII text file – it can be in UTF-8 as well.
3. Examples of information that may be included in the header file are as follows:
    a. Description of the power system prior to disturbance
    b. Name of the station
    c. Identification of the line, transformer, reactor, capacitor, or circuit breaker that experienced the transient
    d. Length of the faulted line
    e. Positive and zero-sequence resistance, reactance, and capacitance
    f. Mutual coupling between parallel lines
    g. Locations and ratings of shunt reactors and series capacitors
    h. Nominal voltage ratings of transformer windings, especially the potential and current transformers
    i. Transformer power ratings and winding connections
    j. Parameters of the system behind the nodes where the data were recorded (equivalent positive- and zero-sequence impedance of the sources)
    k. Description of how the data were obtained, whether it was obtained at a utility substation or by simulating a system condition on a computer program such as an electro-magnetic transient program (EMTP)
    l. Description of the anti-aliasing filters used
    m. Description of analog mimic circuitry
    n. The phase sequencing of the inputs

5.4 DAT File

The data file contains the scaled data values of the sampled event. It includes the sample number, time stamp, and data values of each channel for each sample in the file. The data file may be in ASCII, binary, binary32, or float32 format, which is specified in the corresponding configuration file. Testing to be done as follows:

1. Check for the file extension; it should be **.DAT**.
2. Check the number of lines; there should be one line for each sample. The number of samples is specified in the .CFG line (item 9 in section 5.1); the number of samples for each sampling rate is

summed for all specified sampling rates to determine the total number of samples and hence the number of lines that should be in the .DAT file.

3. Check for the scaled data values.
   a. for an ASCII data file: The ASCII data file should consist of rows and columns. Each row should be divided into TT+2 columns where TT is the total number of channels, analog and status, specified in the .CFG file; the other two columns are for the sample number and time stamp. Each data sample record row should be formatted as follows:

      n, timestamp, A1, A2, …..Ak, D1, D2, ..…Dm

      1. n should be 1-10 integer characters with 1 < n < 9999999999.
      2. timestamp should be 1-13 integer characters.
      3. A1 ….Ak should be 1-13 characters, real or integer and numeric, 3.4028235E38 ? Ak ? -3.4028235E38 and separated by commas. Missing data must be represented by data separators immediately following each other with no spaces (null fields). Each value should fall within the range specified for the channel in the .CFG file (items 6i and 6j in section 5.1).
      4. D1 …..Dm should be 0 or 1, and separated by commas. The last data value should be terminated with carriage return/line feed.

   b. for a binary data file: The binary, binary32 and float32 data files use the same basic structure as that used for the ASCII data files, with the exception that status channel data are compacted. The format is sample number, time stamp, data value for each analog channel, and grouped status channel for each sample in the file. No data separators are used and the data file is a continuous stream of data. Each data sample record shall consist of numeric values arranged as follows:

      n timestamp A_1 A_2 …..A_k S_1 S_2 ..…S_m

      1. **n** should be 4 bytes-long integer number with 00000001 $\leq$ **n** $\leq$ FFFFFFFF (in hexadecimal).
      2. **timestamp** should be 4 bytes-long integer number with 00000000 $\leq$ **n** $\leq$ FFFFFFFE (in hexadecimal). Missing value should be replaced with "FFFFFFFF".
      3. **A_1 ….A_k** should be numeric (real or integer), 2 bytes for binary data files and 4 bytes for binary32 and float32 files. Missing value should be replaced with the corresponding maximum negative value in the field. Each value should fall within the range specified for the channel in the .CFG file (items 6i and 6j in section 5.1).
      4. **S_1 …..S_m** should be 2 bytes integer number with 0000 $\leq$ **Sm** $\leq$ FFFF.

6 Testing of Single File Format (.CFF)

The 2013 COMTRADE standard also provides a single file format with **.CFF** extension. This single file format is merely a collection of the four individual files (.CFG, .INF, .HDR and .DAT) as separate sections. Each section begins with a separator on its own line, which indicates the start of the section. The content and the lines of the .CFF file are defined in order as follows:

1. The first separator indicating the start of the .CFG file section.
2. The entire contents of the configuration section.
3. The second separator indicating the start of the .INF file section.
4. The entire contents of the information section.
5. The third separator indicating the start of the .HDR file section.

6. The entire contents of the header section.
7. The fourth separator indicating the start of the .DAT file section.
8. The entire contents of the data section.
9. The end of the single file shall be indicated using the end of file marker.

Testing of the .CFF file to be done as follows:

1. Check for the first separator (--- file type: CFG ---), it should contain the name of “CFG”.
2. Check for the .CFG file section, the formatting requirement should follow the rules as specified in the section 5.1 for the .CFG file.
3. Check for the second separator (--- file type: INF ---),, it should contain the name of “INF”. Empty lines are allowed before the line of the second separator.
4. Check for the .INF file section, the formatting requirement should follow the rules as specified in the section 5.2 for the .INF file. If the .INF file section is absent, an empty line should be inserted in the place of the .INF file information.
5. Check for the third separator (--- file type: HDR ---),, it should contain the name of “HDR”. Empty lines are allowed before the line of the third separator.
6. Check for the .HDR file section, the formatting requirement should follow the rules as specified in the section 5.3 for the .HDR file. If the .HDR file section is absent, an empty line should be inserted in the place of the .HDR file information.
7. Check for the fourth separator (--- file type: DAT ASCII --- or --- file type: DAT BINARY:xxx ---), it should contain the name of “DAT ASCII” or “DAT BINARY”. Empty lines are allowed before the line of the fourth separator. Please check that total number of bytes in a BINARY DAT file equals to xxx.
8. Check for the .DAT file section, the formatting requirement should follow the rules as specified in the section 5.4 for the .DAT file.
9. Check for the end of file marker, the last line of the file should have the end of file marker.

7 Summary

This report has captured the more common errors that have been found in COMTRADE files throughout the years. Each file (or section) type has been discussed in detail showing on a line-by-line basis the data that should be in a COMTRADE 2013 file. This information should help those looking to validate a COMTRADE 2013 file know where to begin, what to look for, and what tools and methods could be used.

Testing should be automated as much as possible.

The important files in COMTRADE data are the data .DAT and configuration .CFG files.

It is a good idea to view the data in the COMTRADE waveform viewer. Knowing the type of fault, a visual check should be made to make sure the data make sense. As an example, say the file is for A-G fault. Then at the inception of the fault, phase A voltage should be depressed and phase A current should be high. Also in the residual or ground current is recorded it should be the sum of all three phases. Same rule applies to voltage channels. Also to check once the fault is cleared, all the phase voltages either go to zero or return to normal.

Any questions or comments on the testing procedure described in this report may be sent to COMTRADE2013@gmail.com.

Annex A (informative) Bibliography

[1] IEEE Standard Common Format for Transient Data Exchange (COMTRADE) for Power Systems, IEEE C37.111-1991, 1991.

[2] IEEE Standard Common Format for Transient Data Exchange (COMTRADE) for Power Systems, IEEE Std C37.111-1999, 1999.

[3] Electrical relays - Part 24: Common format for transient data exchange (COMTRADE) for power systems, IEC Std. 60255-24 ed1.0, 2001.

[4] Measuring Relays and Protection Equipment - Part 24: Common format for transient data exchange (COMTRADE) for power systems, IEC 60255-24 ed2.0, 2013-04, IEEE Std C37.111™.

[5] IEEE Standard Letter Symbols for Units of Measurement (SI Units, Customary Inch-Pound Units, and Certain Other Units), IEEE Std 260.1-1993.

[6] IEEE Standard Letter Symbols for Quantities Used in Electrical Science and Electrical Engineering, ANSI/IEEE Std 280-1985, Reaffirmed 1997, 2003.

[7] ISO 80000-1:2009 Quantities and units – Part 1: General

Annex B (informative) Information File/Section Details

B.1 Information file structure

The information file is divided into sections. Each section consists of a header line followed by a number of entry lines. There is no limit to the number of sections but there shall be at least one section per file. No data shall reside outside of a section. Each section is identified by a unique section header line. All data belong to the nearest section header above it in the file.

Generically the structure is as follows:

Public Record Information Section Header (information relating to the whole record)
    Publicly-Defined Record Information Entry Lines
Public Event Information Section Header (information relating to a particular channel and sample in the record)
    Publicly-Defined Event Information Entry Lines
Public File Description Section Header (information equivalent to .CFG file information relating to the whole record)
    Publicly-Defined File Description Entry Lines
Public Analog Channel #1 Section Header (information equivalent to .CFG file information relating to the first analog channel in the record)
    Publicly-Defined Analog Channel Entry Lines
Public Analog Channel #n Section Header (information relating to the next analog channel in the record, with a new section for each channel, up the number of analog channels in the record)
    Publicly-Defined Analog Channel Entry Lines
Public Status Channel #1 Section Header (information relating to the first status channel in the record)
    Publicly-Defined Status Channel Entry Lines
Public Status Channel #n Section Header (information relating to the next status channel in the record, with a new section for each channel, up the number of status channels in the record)
    Publicly-Defined Status Channel Entry Lines
Private Information Header
    Privately-Defined Record Information Entry Lines
Private Information Header
    Privately-Defined Record Information Entry Lines

B.2 Examples of header naming

Public sections:
Examples:

Acceptable:
[Public File_Description] <CR/LF>

Unacceptable:
[Public DataSource]<CR/LF> (Leading space)
[DataSource Public]<CR/LF> (Must begin with word Public)

Private sections:
Examples:

Acceptable:
[Company1 Input Ranges]<CR/LF>

[Company2 IsolatorType] <CR/LF>

Unacceptable:

[Company Name Input Ranges] <CR/LF>	(Spaces not allowed in owner identifier)
[12] <CR/LF>	(Starts with number)
{Bad Section}<CR/LF>	(Wrong bracket style)
[Bad Section<CR/LF>	(Missing bracket)
[Bad Section] Extra Data=Not Allowed<CR/LF>	(Extra text or entries on line after closing bracket)

B.3 Examples of entry line

An entry line must start with one word 3–32 characters long followed by an equal (=) sign. The first word is the “Entry Name.” The entry name is a description of the function of the value string that follows. It is analogous to the name of a variable or constant in many programming languages. The entry name shall be meaningful when read in conjunction with the section name. The entry name need not be fully descriptive. The entry name can contain any printable characters with ASCII values between 33 and 127 decimal. The line shall be terminated with a <CR/LF>.

Examples:

Acceptable:

[Public File Description] <CR/LF>
Recording_Device_ID=Unit 123<CR/LF>

[Company2 Calibration] <CR/LF>
Ch1=2044.5, –7, 1<CR/LF>
Ch2=2046.2, 5.3, 1<CR/LF>
Ch3=2042.0, –0.4, –1<CR/LF>

Unacceptable:

[Company3 Calibration] <CR/LF>	
cl33421thvlst=2044.5,–7,1,2046.2,5.3,1,2042.0,–0.4,–1<CR/LF>	(Entry name not meaningful)
Ch 1= 2044.5, –7, 1<CR/LF>	(Extra spaces)
[Company3 Device Type] <CR/LF>	(No space between last entry in previous section)

B.4 Comment lines

An entry line prefixed with a semicolon is considered a comment line. Such lines are to be skipped by file reading algorithms and are used for comments or to comment out certain entries. When section headings are commented out, all entry lines in that section shall also be commented out. Failing to comment out the entry lines in a section where the heading has been commented out would cause any uncommented entry lines in that section to fall under the previous section heading.

Examples:

Acceptable:

[Company2 Calibration] <CR/LF>
; Sequence is gain, offset, polarity<CR/LF>
Ch1=2044.5, –7, 1<CR/LF>

Ch2=2046.2, 5.3, 1<CR/LF>
;Channel 2 replaced 7/16/95<CR/LF>
Ch3=2042.0, –0.4, –1<CR/LF>

Unacceptable:

;[Company3 Calibration] <CR/LF>	(Section heading commented out leaving orphan data)
Ch 1 = 2044.5, –7, 1<CR/LF>	(Extra spaces)
;This recorder uses 8 bit data and has 64 channels, test points on the card are high impedance and not galvanically isolated. <CR/LF>	(excessive and wrongly-placed documentation)

B.5 Value string

The value string is defined as all characters on an entry line from the equal sign to the end-of-line sequence. Value strings can contain one data item or several data items. Multiple data items are separated by commas. Numeric values shall begin immediately after the equal sign or comma delimiter with no leading space. Text strings that include a space after the equal sign or comma delimiter shall include the space as part of the value. For public sections, this information is specified in the standard. For private sections, the data type, format, and number of items per entry line are defined by the user.

B.6 Publicly-defined sections

The 2013 COMTRADE standard specifies some public section headers and entry lines. If a publicly-defined section header is included, all of the defined entry lines for the section shall be included in the order specified. Up to 2013, the publicly defined sections included:

a) Public Record Information Section;
b) Public Event Information Section;
c) Public File Description Section;
d) Public Analog Channel #n Section; and
e) Public Status Channel #n Section.

B.7 Sample .INF file

[Public Record_Information]<CR/LF>
Source=COMwriter, V1.1<CR/LF>
Record_Information=Fault, AG, Trip,Transmission Line<CR/LF>
Location=189.2, miles<CR/LF>
max_current=3405.5<CR/LF>
min_current=–3087.2<CR/LF>
max_voltage=208.6<CR/LF>
min_voltage=–206.4<CR/LF>
EventNoteCount=2<CR/LF>
<CR/LF>
[Public Event_Information_#1] <CR/LF>
Channel_number=2<CR/LF>
max_value=204.5<CR/LF>
min_value=–205.1<CR/LF>
max_sample_number=168<CR/LF>
min_sample_number=15<CR/LF>

```
Sample_number_Text_#1=168,Transient on reclose<CR/LF>
Sample_number_Text_#2=15,Minimum during normal load <CR/LF>
<CR/LF>
[Public Event_Information_#2] <CR/LF>
Channel_number=1<CR/LF>
max_value=206.5<CR/LF>
min_value=205.1<CR/LF>
max_sample_number=159<CR/LF>
min_sample_number=9<CR/LF>
Sample_number_Text_#1=159,Transient on reclose<CR/LF>
Sample_number_Text_#2=9,Minimum during normal load <CR/LF>
<CR/LF>
[Public File_Description] <CR/LF>
Station_Name=Condie<CR/LF>
Recording_Device_ID=518<CR/LF>
Revision_Year=1999<CR/LF>
Total_Channel_Count=12<CR/LF>
Analog_Channel_Count=6<CR/LF>
Status_Channel_Count=6<CR/LF>
Line_Frequency=60<CR/LF>
Sample_Rate_Count=1<CR/LF>
Sample_Rate_#1=6000.000<CR/LF>
End_Sample_Rate_#1=885<CR/LF>
File_Start_Time=11/07/95,17:38:26.663700 <CR/LF>
Trigger_Time=11/07/95,17:38:26.687500 <CR/LF>
File_Type=ASCII <CR/LF>
Time_Multiplier=1<CR/LF>
<CR/LF>
[Public Analog_Channel_#1] <CR/LF>
Channel_ID=Popular Va-g<CR/LF>
Phase_ID=<CR/LF>
Monitored_Component=<CR/LF>
Channel_Units=kV<CR/LF>
Channel_Multiplier=0.14462<CR/LF>
Channel_Offset=0.0000000000<CR/LF>
Channel_Skew=0<CR/LF>
Range_Minimum_Limit_Value=–2048<CR/LF>
Range_Maximum_Limit_Value=2048<CR/LF>
Channel_Ratio_Primary =2000<CR/LF>
Channel_Ratio_Secondary=1<CR/LF>
Data_Primary_Secondary=P<CR/LF>
<CR/LF>
[Public Status_Channel_#1] <CR/LF>
Channel_ID=Va over<CR/LF>
Phase_ID=<CR/LF>
Monitored_Component=<CR/LF>
Normal_State=0<CR/LF>
<CR/LF>
[Company1 event_rec] <CR/LF>
```

recorder_type=1<CR/LF>
trig_set=0,0,0,0,6048,6272,0,0,0,0,0,0,0,0,0,0<CR/LF>
ch_type=1,1,1,1,1,1,1,0,0<CR/LF>
<CR/LF>
[Company1 analog_rec_#1] <CR/LF>
op_limit=15<CR/LF>
trg_over_val=f<CR/LF>
trg_under_val=f<CR/LF>
trg_roc=f<CR/LF>
inverted=0<CR/LF>
<CR/LF>