Decoding and Conversion of TFD Logging Data Format Based on Java

Mingxing Liang *

State Key Laboratory of Coal Resources and Safe
Mining, China University of Mining and Technology
(Beijing), Beijing, China
e-mail: m.liang.cn@gmail.com

Peipei Zhang

College of Geoscience and Surveying Engineering,
China University of Mining and Technology (Beijing),
Beijing, China
e-mail: zhpxfcy@163.com
* Corresponding Author

Abstract— Matrix logging system is widely used in coal-bed methane (CBM) reservoir. TFD is the extension of Matrix output file which is a structured format comprised a mix of ASCII and binary. WellCAD is the only logging software which can be used because of the confidentiality format, so that the processing and interpretation results of CBM reservoir are less effective. This paper decodes the TFD format and analyzes the storage pattern. Based on this, Java programming language is used to develop a format conversion program. The application can accurately convert the TFD format data to LAS (Log ASCII Standard), which can be loaded into the majority of well logging software platform. Result shows that the description of DAT structure and outcome of DAT file conversation are completely accurate. With the help of those powerful software, the efficiency of CBM well logging data processing and interpretation will be extremely improved.

Keywords-TFD; LAS; Java; data structure; well logging

I. INTRODUCTION

Mount Sopris Instruments Company founded in 1951 is a recognized world pioneer in small diameter geophysical logging instrument for groundwater, mining, geotechnical and other applications [1]. In the mid-1980s, the company reached a series of agreements of technology transfer and cooperation with relevant coal authorities in China, which vigorously promoted the rapid development of Chinese coal logging technology. Mount Sopris Matrix is the most versatile data acquisition system which can be used with all Mount Sopris probes, past and present, as well as many third party tools. This versatility, combined with its compact size and dependable performance, has made the Matrix the most popular logger in the coalfield and coal bed methane (CBM) industry of China.

LoggerSuite is the entire software package for operating the data acquisition hardware such as Matrix, Abox, Bbox, and AltLogger Logging systems [2]. Logger, a part of LoggerSuite, is used for acquiring logging data. This software takes TFD file as the database. TFD format is a structured binary file, which can only be opened with WellCAD logging software [3]. WellCAD is PC based composite log package, which combines comprehensive graphic editing mechanisms and data processing tools, and requires a dongle to run. In addition, WellCAD lacks appropriate methods of log interpretation in

unconventional reservoirs, especially in coalbed methane reservoir. Therefore, it is essential to decode and convert the TFD data format.

At present, the engineers use the highest frequency of logging software including Techlog, Geolog, Forward, CIFLog, Lead, etc. These systems do not only have excellent display, editing and analysis capabilities for well data, but also support kinds of well logging data format except the TFD. The investigation finds that LAS format is the most widely used well logging format, and also can be loaded by most of well logging software platforms. LAS format created by Canadian Well Logging Society is a structured ASCII file containing header information and log curve data. This paper mainly analyses the data structure of TFD and LAS file, and develops a format-conversion program, which can assist various well logging software platforms to load and process TFD data.

II. DATA STRUCTURE ANAYSIS

Well logging data formats are specified data structures formatted in a certain way for acquiring, ex-changing, and storing logging data. Main recorded object of the logging data format is the curve data, which is composed by a set of index and a collection of values. Index is commonly depth or time, and value can be a single value or multiple values. Curve data is mainly divided into three types: 1) routine logging curve (one-dimensional curve); 2) imaging logging curve (two-dimensional curve); 3) discrete logging curve (discrete data). The TFD and LAS 2.0 format described in this paper can only store routine logging curves, whose index is equal-interval and the value of each index is a single value. A routine curve L can be expressed,

$$L = \left\{ \left(C_i, K_i \mid i \in [1,2,3 \cdots n] \right) \right\} \quad (1)$$

$$C_i = C_1 + I_s \times (i - 1) \quad (2)$$

where Ci is the index (depth or time) of the i-th sampling point; Ki is the value of the i-th sampling point; and Is is the sampling interval.

A. TFD Data Structure Analysis

TFD file is a structured mix of ASCII and binary formats (Fig. 1). Each TFD file begins with a file header record and is followed by one or more data packets. The file header is composed of ASCII and binary formats, but the data packet is only comprised of binary format.
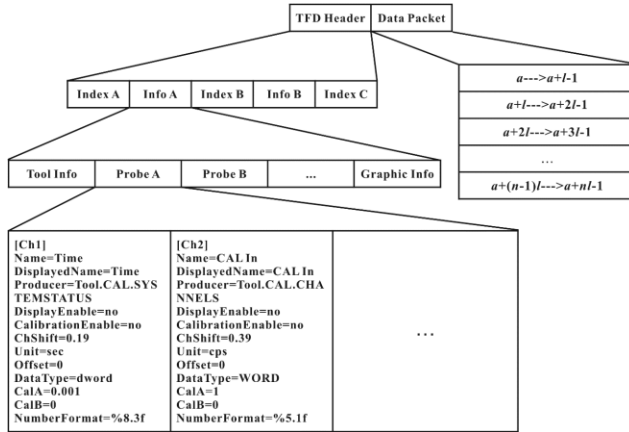
Figure 1. Structure of TFD file

TFD Header consists of Index A, Info A, Index B, Info B, and Index C. These three Indexes are binary and the byte length are fixed. Index A is 116 bytes in size. It mainly records the starting position of Index B and the file creation date. Index B is 234 bytes long and records the starting position of Index C. Index C is 150 bytes and records the byte length of each data packet (Table 1). What the Indexs record are critical for decoding TFD file because it influence the arrangement location of data blocks.

TABLE I. DESCRIPTION OF MAJOR RECORD IN INDEXES

Part	Data Type	Starting position	Bytes	Description
Index A	Short	0	2	The starting position of Index B: I_1
	Short	42	12	Date: YYYY / MM / DD / HH / MM / SS
Index B	Short	I_1+8	2	The starting position of Index C: I_2
Index C	Short	I_2+138	2	The byte length of each data packet: l

Info A and Info B are ASCII formats, and the bytes length of them are variable. Info A expounds the information of combined probes, the curves' information of each linked probe and the logging plot information. Info B mainly records the configuration information of Logger system. Curves' information elaborate the data type, number format and other important parameters of data storage. Table 2 shows curves' information of the natural gamma measurement probe.

TABLE II. DESCRIPTION OF CURVES FOR QL40-GAM PROBE

Name	Data Type	Cal A	Cal B	Number Format	Unit
Time	Int	0.001	0	%8.3f	Sec
TCPU	Short	0.217226	-61.11	%5.1f	℃
EHT	Short	-3.65885	0	%5.0f	Volts
Count	Int	1	0	%5.0f	
GR		1	0	%6.2f	CPS

This probe can obtain five curves, which have different data types for storage. According to the description, the format of log curve data in TFD file is integer or short integer, however, Number Specifier defined the format of final log curve data in WellCAD is float. By contrasting the TFD and WellCAD data, the conclusion has to be come to that these data satisfy the (3),

$$D_W = C_A \times D_T + C_B \quad (3)$$

where DW is the final log curve data in WellCAD, DT is the original log curve data in TFD file, CA and CB presented Cal A and Cal B separately, which are the parameters described in the curves' information.

Data package part starts from the end of Index C, and each block represented one point (depth or time) of data. The byte length of each block has been described in Index C.

LAS Data Structure Analysis

The LAS format was invented by the Canadian Well Logging Society in 1989 to standardize the organization of digital log curve information for personal computer users [4]-[10]. It did this very successfully and the standard became popular worldwide. LAS file contains for kinds of flags, which represents different information. The implications of these flags are summarized in Table 3.

TABLE III. DEFINITION OF FLAGS IN LAS 2.0

Flag	Description
~	The beginning of each section
#	The beginning of each comment line
.	Line delimiter, before the units and after the mnemonic
:	Line delimiter, before the description and after the data

Each LAS file consists of sections. Sections begin with a header line defined as beginning with the "~" (tilde character) when it occurs as the first non-space character on a line. The character immediately follows the tilde character defines the section with the remainder of the line being ignored. The characters "V", "W", "C", "P", "O", and "A" reserved in the LAS represent the meanings of sections (Table 4).

TABLE IV. DEFINITION OF FIELDS IN LAS 2.0

Field	Description	Comment
~V	Version Information	Mandatory and must appear as the first section in the file
~W	Well Information	Mandatory
~C	Curve Information	Mandatory
~P	Parameter Information	Optional. It defines the input values of various parameters relating to this well.
~O	Other Information	Optional. It is intended as a remarks or comments section.
~A	ASCII Log Data	This data section will always be the last section in a file.

The sections defined by the LAS 2.0 standard are limited to one occurrence per file. Customer defined sections are permitted but must be located after the first section (~V) and before the last section (~A). Fig. 2 is an

example of a LAS 2.0 file with minimal header information in unwrapped mode.

```
~V
VERS.           2.0    :  CWLS LAS-VERSION 2.0
WRAP.           NO     :  One line per depth step
~W
STRT.M                635.0000      :START DEPTH
STOP.M                634.8750      :STOP DEPTH
STEP.M                -0.1250       :STEP
NULL.                 -999.25       :NULL VALUE
COMP.     ANY OIL COMPANY INC.   :COMPANY
WELL.     ANY ET AL 12-34-12-34  :WELL
FLD .     WILDCAT                :FIELD
LOC .     12-34-12-34W5M         :LOCATION
PROV.     ALBERTA                :PROVINCE
SRVC.     LOGGING COMPANY INC.   :SERVICE COMPANY
DATE.     13-DEC-86              :LOG DATE
UWI .     100123401234W500       :UNIQUE WELL ID
LIC .     23412                  :ERCB LICENCE NUMB
~C
DEPT    .M       :   DEPTH
RHOB    .K/M3    :   BULK DENSITY
NPHI    .VOL/VOL :   NEUTRON POROSITY - SANDSTONE
MSFL    .OHMM    :   Rxo RESISTIVITY
SFLA    .OHMM    :   SHALLOW RESISTIVITY
ILM     .OHMM    :   MEDIUM RESISTIVITY
ILD     .OHMM    :   DEEP RESISTIVITY
SP      .MV      :   SPONTANEOUS POTENTIAL
~A
635.0000  2256.000  0.403  22.078  22.078  20.343  23.666  123.4
634.8750  2256.000  0.403  22.078  22.078  20.343  23.666  123.4
634.7500  2256.000  0.403  22.078  22.078  20.343  23.666  123.4
634.6250  2256.000  0.403  22.078  22.078  20.343  23.666  123.4
```

Figure 2. LAS 2.0 file with minimal header information

III. PROGRAM DESIGN AND DEVELOPMENT

After being discussed, the structures of TFD and LAS file are figured out, which are two distinct formats. In order to develop a format converter, the paper establishes three program packages: "tfdreader", "laswriter" and "tfd2lasconverter". Among them, "tfdreader" package and "laswriter" package are primary, and the "tfdreader" package includes four classes. As can be seen from Fig. 3, the "TFDReader" class realized four classes achieve TFD reading function. The "TFDIndexInfo" class achieves reading function of TFD Index information. The "TFDCurveInfo" class achieves reading function of curve information. The "TFDCurveData" class achieves reading function of curve data.
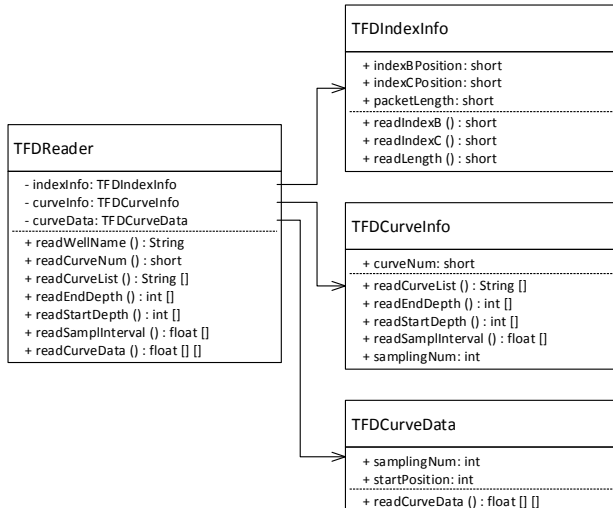
Figure 3. TFD class diagram of "tfdreader" package.

The "laswriter" package includes seven classes (Fig .4). The "LASWriter" realized six classes achieves LAS writing function. These six classes achieve writing function of LAS version information, well information, curve information, parameter information, other information and curve data separately.
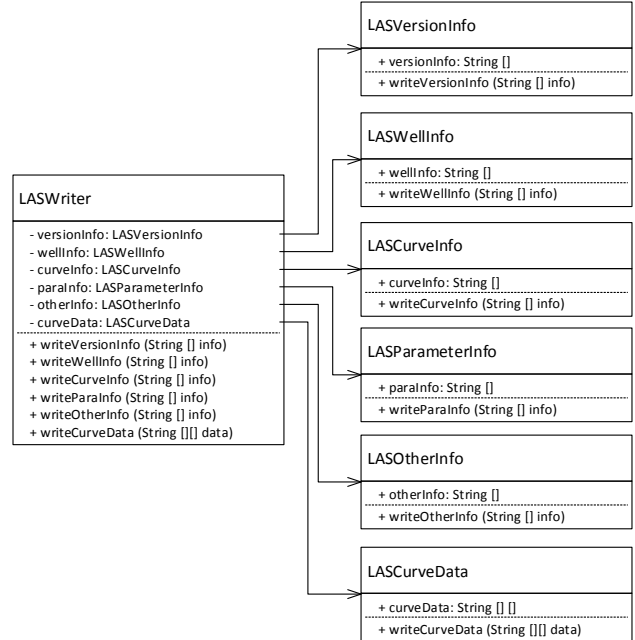
Figure 4. UML class diagram of "laswriter" package.

After finishing the two decoding program development, the paper designs an interactive interface for TFD data format converter (Fig. 5). The interface contains five tabbed panes including well information display panel, curve information display panel, curve data display panel, layer data display panel and log plot display panel. When users open a TFD file, the information and data of the corresponding well will be loaded and displayed.
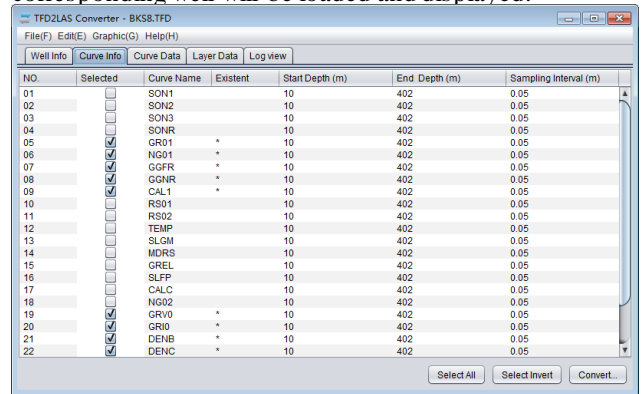
Figure 5. Application's interface of the converter.

IV. APPLICATION AND VERIFICATION

Currently, the most often used well logging software platform include Techlog, Geolog, WellCAD, etc. Techlog, developed by Schlumberger, is a Windows based software platform planned to integrate all the wellbore information. The Techlog platform allows users to have a highly interactive and intuitive way to load, observe, and analysis their data. Geolog produced by Paradigm combines the

system's technological superiority with a modern, user-friendly interface that ensures improved efficiency and usability. WellCAD is PC based composite log package, which combines integrated graphic editing mechanisms and data processing tools. These platforms all support the LAS data format.

In the software above, Techlog has more powerful functions. Hence, the paper chooses Techlog to contrasting and confirming with the TFD Converter. In this article, the BKS8 well logging data (TFD format) is taken as an example to express the procedures of data conversation and log-plot drawing:

1) Load the TFD file by TFD Converter. Nine curves are drew in the Log view tabbed pane (Fig. 6).

Figure 6. The log-plot of well BKS8 in TFD Converter.

2) Select the valid curves to convert into LAS file.
3) Load the LAS file into Techlog platform and use the LogView function to draw a log graph (Fig. 7).
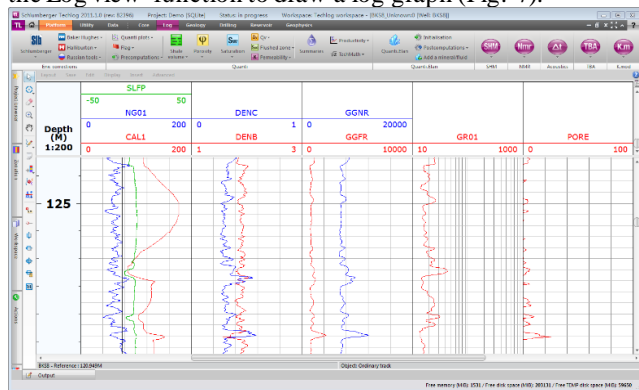

Figure 7. The log-plot of well BKS8 in Techlog platform.

Comparing Fig. 6 with Fig. 7, the curve shape of the different software display is unanimous, which illustrates the correctness and effectiveness of the procedures.

V. CONCLUSION

A DAT2LAS converter is presented for converting a DAT binary file into a LAS text file. Although the specification of LAS file format is open, DAT files must be decoded because binary-encoded form are used to store well logging data in DAT format. This paper expounds the structure of DAT file and develops a converter program. Result shows that the description of DAT structure and outcome of DAT file conversation are completely accurate, which assists the processing and interpretation of coalfield well logging data in a significant way.

REFERENCES

[1] John R. Stowell, and James J. Lococo, "Borehole radiometrics - Present trends & future direction," Environmental and Engineering Geophysical Society - 21st Symposium on the Application of Geophysics to Engineering and Environmental Problems, vol. 1, 2008, pp. 701-710.

[2] John R. Stowell, "Optimizing the measurement of fluid low in boreholes - A review of current hardware and applications," 26th Symposium on the Application of Geophysics to Engineering and Environmental Problems 2013, SAGEEP 2013, 2013, pp. 589-600.

[3] John R. Stowell, and Lia M. Martinez, "The application of Monte Carlo simulation to borehole gamma-gamma density and spectral gamma calibrations," 27th Symposium on the Application of Geophysics to Engineering and Environmental Problems 2014, SAGEEP 2014, 2014, pp. 295-302.

[4] R.Y. Elphick, "Using the LAS format-Part 1," Geobyte, vol. 6, iss. 6, 1991, pp. 44-61.

[5] R.Y. Elphick, "Using the LAS format-Part 2," Geobyte, vol. 7, iss. 1, 1992, pp. 22-33.

[6] R.Y. Elphick, "Using the LAS format-Summations," Geobyte, vol. 7, iss. 2, 1992, pp. 38-52.

[7] R.Y. Elphick, "Log analysis within the LAS format," Geobyte, vol. 34, iss. 2, 1993, pp. 67-90.

[8] C. Struyk, R. Bishop, and D. Fortune, et al., "LAS, a flop-py disk standard for log data," Geobyte, vol. 30, iss. 5, 1989, pp. 395-396.

[9] C. Struyk, G. Gray, and G. Peterson, et al., "New features of LAS version 2.0; the floppy disk standard for log data," The Log Analyst, vol. 34, iss. 2, 1993, pp. 60-66.

[10] S. I. Zkaya, "An EXCEL macro for importing log ASCII standard (LAS) files into EXCEL worksheets," Computers and Geosciences, vol. 22, iss. 1, 1996, pp. 75-80.