SANDIA REPORT

SAND201X-XXXX
Unlimited Release
Printed August 2017

PCF File Format

Gregory G. Thoreson

Prepared by
Sandia National Laboratories
Albuquerque, New Mexico 87185 and Livermore, California 94550

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

Issued by Sandia National Laboratories, operated for the United States Department of Energy by National Technology and Engineering Solutions of Sandia, LLC.

Available to DOE and DOE contractors from
U.S. Department of Energy
Office of Scientific and Technical Information
P.O. Box 62
Oak Ridge, TN 37831

Telephone: (865) 576-8401
Facsimile: (865) 576-5728
E-Mail: reports@osti.gov
Online ordering: http://www.osti.gov/scitech

Available to the public from
U.S. Department of Commerce
National Technical Information Service
5301 Shawnee Rd
Alexandria, VA 22312

Telephone: (800) 553-6847
Facsimile: (703) 605-6900
E-Mail: orders@ntis.gov
Online order: http://www.ntis.gov/search

SAND2017-9107
Printed August 2017
Unlimited Release

PCF File Format

Gregory G. Thoreson
WMD Detection & Threat Analysis
Sandia National Laboratories
P. O. Box 5800
Albuquerque, New Mexico 87185-MS0782

Abstract

PCF files are binary files designed to contain gamma spectra and neutron count rates from radiation sensors. It is the native format for the GAmma Detector Response and Analysis Software (GADRAS) package [1]. It can contain multiple spectra and information about each spectrum such as energy calibration. This document outlines the format of the file that would allow one to write a computer program to parse and write such files.

TABLE OF CONTENTS

TABLES

NOMENCLATURE

Abbreviation	Definition
Abbreviation	Definition
GADRAS	GAmma Detector Response and Analysis Software
DHS	Department of Homeland Security
DNDO	Domestic Nuclear Detection Office

1. INTRODUCTION

The PCF binary format allows the start of each spectrum in the file to be aligned at a known byte location in the file, which in turn allows programs to extract a single spectrum without parsing the entire file into memory. Therefore, the contents of the file are described in this document relative to *byte offsets*. The start of the file is byte offset zero.

2. FILE HEADER

The first number in the file header is a 2-byte signed integer that is the number of "records" per spectrum ($NRPS$). Each spectrum can have a different number of channels. $NRPS$ determines the maximum number of channels allowed so they can all be byte-aligned. The maximum number of channels allowed for any record is given by

$$C_{max} = 64(NRPS - 1)$$

The second value in the file is a three-character string. If it is equal to "DHS", the header is longer and contains more detailed information about a vehicle occupancy in a radiation detection portal, typically of interest to the Department of Homeland Security (DHS). Otherwise, the initial header is much shorter.

2.1. Non-DHS Version File Header

The short file header data is summarized in Table 1. The meaning of the energy calibration coefficients can be found in reference [1]. The energy calibration coefficients found in the file header are typically not used, as each spectrum can have its own energy calibration coefficients.

Table 1. Non-DHS Version File Header

Byte Offset	Data Type	Meaning
0	2-byte signed integer	Number of records per spectrum (NRPS)
2	3-byte character array	Version
5	4-byte character array	Energy calibration label (unused)
9	4-byte floating point	Energy calibration offset (keV)
13	4-byte floating point	Energy calibration gain (full-range fraction)
17	4-byte floating point	Energy calibration quadratic term (full-range fraction)
21	4-byte floating point	Energy calibration cubic term (full-range fraction)
21	4-byte floating point	Energy calibration low-energy (full-range fraction)

2.2. DHS Version File Header

The extended file header data format is summarized in Table 2.

Table 2. DHS Version File Header

Byte Offset	Data Type	Meaning
0	2-byte signed integer	Number of records per spectrum (NRPS)
2	3-byte character array	Version = "DHS"
5	7-byte character array	File last modified date hash
12	36-byte character array	UUID
48	16-byte character array	Inspection
64	2-byte signed integer	Lane number
66	26-byte character array	Measurement remark
92	28-byte character array	Instrument type
120	28-byte character array	Manufacturer
148	18-byte character array	Instrument model
166	18-byte character array	Instrument ID
184	20-byte character array	Item description
204	16-byte character array	Measurement location name
220	16-byte character array	Measurement coordinates
236	2-byte signed integer	Item to detector distance
238	2-byte signed integer	Occupancy number
240	16-byte character array	Cargo type

2.3. Deviation Pairs and Spectral Start Index

After the initial 256 bytes, this ends the nominal header information. At the 256 byte offset mark there is an optional deviation pair indicator string.

At this offset, the parsing program must parse for a string to check if it is equal to "DeviationPairsInFile" (20 bytes), or "DeviationPairsInFileCompressed" (30 bytes). If it is "DeviationPairsInFile" or "DeviationPairsInFileCompressed", then the spectral record start index ($SRSI$), which determines where spectral information begins in the file, is 83. If this string is anything else, $SRSI$ is 2.

2.3.1. Normal Deviation Pairs

If the string at byte offset 256 is "DeviationPairsInFile" and not "DeviationPairsInFileCompressed", then there are 5,120 4-byte floating point values (20,480 bytes total) starting at an offset of 512 bytes in the file.

Each deviation pair is an energy (usually a photopeak) in keV, and the deviation/offset in keV. Each detector can have 20 deviation pairs. There can be 128 different detectors defined by their column index, panel index, and MCA index. The pseudo-code for parsing order is:

loop over columns (2)
    loop over panels (8)

loop over MCAs (8)
loop over deviation pairs (20)
parse energy (4-byte floating point)
parse offset (4-byte floating point)

2.3.2. Compressed Deviation Pairs

If the string at byte offset 256 is "DeviationPairsInFileCompressed" then this allows 4 columns instead of the normal 2 columns by compressing the data into 2-byte integers. There are 10,240 2-byte signed integers, which is the same 20,480 bytes total and occupies the same space in file as normal deviation pairs.

The pseudo-code for parsing the compressed values is similar:

loop over columns (4)
loop over panels (8)
loop over MCAs (8)
loop over deviation pairs (20)
parse compressed energy (2-byte signed integer)
parse compressed offset (2-byte signed integer)

After parsing the values, the energies should be converted to a floating point value. The offsets should be divided by 10 and converted to a floating point value.

2.4. Total Number of Spectral Records in File

The number of spectral records, N, can now be determined from the total size of the file in bytes, F, and the other two parsed values $SRSI$ and $NRPS$:

$$N = \frac{F - 256(SRSI - 2)}{256 * NRPS}$$

NOTE: If the last spectral record has less channels than the maximum number allowed, then when writing a PCF it is imperative that you pad the spectra record data with zeroes to make the file size correct. Otherwise, some parsers will not see the last record.

3. SPECTRAL DATA

3.1. Header

To parse the j^{th} spectral record header in the file (starting at 1), the byte offset equation is

$$256(SRSI + NRPS(j - 1) - 1)$$

Relative to this offset, the spectral header information is summarized in Table 3.

Table 3. Spectral Record Header Data Format

Byte Offset	Data Type	Meaning
0	180-byte character array	Compressed text buffer
180	23-byte character array	Date/time in VAX format (DD-MMM-YYYY HH:MM:SS.SS)
203	1-byte character	Tag
204	4-byte floating point	Live time (seconds)
208	4-byte floating point	Total time / real time (seconds)
212	4-byte floating point	unused
216	4-byte floating point	unused
220	4-byte floating point	unused
224	4-byte floating point	Energy calibration offset (keV)
228	4-byte floating point	Energy calibration gain (full-range fraction)
232	4-byte floating point	Energy calibration quadratic term (full-range fraction)
236	4-byte floating point	Energy calibration cubic term (full-range fraction)
240	4-byte floating point	Energy calibration low-energy term (full-range fraction)
244	4-byte floating point	Occupancy flag (0 = unoccupied, 1 = occupied)
248	4-byte floating point	Total Neutron Counts*
252	4-byte signed integer	Number of channels in spectrum

** Typically, the neutron counts are for an attached neutron detector which does not share the same dead-time as the gamma detector (e.g. a He-3 tube). Therefore, most applications divide the total neutron counts by the total time and not the live time to get the count rate.*

The compressed text buffer containing the title, description, and source is parsed into three separate strings. If the first character in the buffer is the extended ASCII character 255, there should be two more instances of the delimiter to divide the title, description, and source in that order. If there is no delimiter, then characters 1 through 60 are the title, 61 through 120 is the description, and 121 through 180 are the source.

3.2. Channel Data

To parse the j^{th} spectral record channel data in the file (starting at 1), the byte offset equation is

$$256(SRSI + NRPS(j - 1))$$

At this offset, there is an array of 4-byte floating point values corresponding to the number of counts in each channel. The total number of floating point value is C_{max}. The number of useful values is determined by the number of channels specified in the spectral header information.

NOTE: Because each record can have fewer channels than the max allowed, there may be values after the last channel which are garbage, zeroes, or other numbers to be ignored until the start of the next spectral record header.

3.3. Title Conventions

This file format does not have support for ancillary information such as measurement distance or detector name, so the title field in each spectral record is used with keywords to store this information.

3.3.1. Detector Name

The detector name can be specified in the title by using the keyword "Det=XXX" where XXX is a string that follows the Domestic Nuclear Detection Office (DNDO) naming convention (e.g. Aa1, Ba1, Aa1N). The first three letters indicate the panel number, column number, and MCA number, respectively. For example, detector "Da2" is the fourth panel, first column, and second MCA. Appending an "N" indicates it is a neutron detector and not a gamma detector.

This keyword is especially useful when deviation pairs for multiple detectors are specified by the panel/column/MCA indices, or when search data is stored in a PCF from multiple detectors and a user wants to extract signals from individual detectors.

3.3.2. Distance

The distance can be specified using "@ XX yy" or "D=XX yy" where "XX" is a number and "yy" is the unit.

3.3.3. Height

The height can be specified using "H=XX yy" where "XX" is a number and "yy" is the unit"

REFERENCES

1. S.M. Horne et. al., *GADRAS-DRF 18.6 User's Manual*, SAND2016-4345, Sandia National Laboratories, Albuquerque, NM, May 2016.

DISTRIBUTION

1	MS0782	Steven M. Horne	6634 (electronic copy)
1	MS0782	Gregory G. Thoreson	6634 (electronic copy)
1	MS0782	Lisa A. Theisen	6634 (electronic copy)
1	MS0782	Dean J. Mitchell	6634 (electronic copy)
1	MS0782	Michael W. Enghauser	6634 (electronic copy)
1	MS0782	Sean O'Brien	6634 (electronic copy)
1	MS0782	Lee T. Harding	6634 (electronic copy)
1	MS0899	Technical Library	9536 (electronic copy)

Sandia National Laboratories