LAWRENCE LIVERMORE NATIONAL LABORATORY

LLNL-TR-770400

Software Requirement Document: Benchmarking Algorithms for RadioNuclide Identification (BARNI)

M. Monterial, K. E. Nelson, S. E. Labov, S. Sangiorgio

March 25, 2019

This work performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344.

NSDD Partnership
Working Together to Prevent Nuclear Trafficking
The Office of Nuclear Smuggling Detection and Deterrence

Office of Nuclear Smuggling Detection and Deterrence
Science and Engineering Team

Software Requirement Document:

Benchmarking Algorithm for RadioNuclide Identification (BARNI)

05/08/2019

Office of Defense Nuclear Nonproliferation
National Nuclear Security Administration
U.S. Department of Energy

National Nuclear Security Administration

Lawrence Livermore National Laboratory

This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under contract DE-AC52-07NA27344.

LLNL-TR-770400

This page intentionally left blank.

Executive Summary

This document provides the requirements for the Benchmark Algorithm for Radio-Nuclide Identification (BARNI) software, an open-source tool for comparing radionuclide identification algorithm performance. It specifies the requirements for the user interface, maintainability, documentation and testing. In addition, this document outlines the core functionality of the product, which will align with the needs of the Office of Nuclear Smuggling Detection and Deterrence (NSDD). Specifically, these include interoperability with RASE, and compatibility with the FLIR R400.

Table of Contents

Acronyms and Abbreviations

AT	Anti-Terrorism
BARNI	Benchmark Algorithm for RadioNuclide Identification
CCL	Commerce Control List
DOE	Department of Energy
ECCN	Export Control Classification Number
NSDD	Nuclear Smuggling Detection and Deterrence
RASE	Replicate Assessment of Spectroscopic Equipment
RIID	Radioisotope Identification Devices

Definitions

shall	Used to indicate an obligatory requirement
should	Used to indicate an optional requirement

This page intentionally left blank.

1 Introduction

The Department of Energy (DOE) Office of Nuclear Smuggling Detection and Deterrence (NSDD) seeks to improve the performance of RadioIsotope Identification Devices (RIIDs). RIIDs enable users to identifying threat nuclides in applications such as secondary search, facility inspection, and emergency response. As part of this effort NSDD has developed Replicative Assessment of Spectroscopic Equipment (RASE), an open-source tool which generates synthetic spectra and analyzes them with vendor's radionuclide identification software in order to evaluate performance.

RASE lacks a radionuclide identification algorithm to perform analysis independent of vendor provided algorithms. An independent radionuclide identification algorithm could serve as a common benchmark against which all other vendor algorithms could be compared against. In addition, the wider community could benefit from a tool which could provide a common measure of performance. Performance metrics used to test algorithms found in literature cannot be readily compared against one another, but a common algorithm could provide a more objective measure. In addition, such an algorithm would provide a meaningful "floor" for identification performance, and further spur innovation in the research community.

Due to various export control, intellectual property and commercial considerations, no open-source tool exists that could perform automatic radionuclide identification on spectra from a diverse set of instruments. This has spurred NSDD to develop the Benchmark Algorithm for Radio-Nuclide Identification (BARNI), an open-source radionuclide identification companion tool for RASE. Unlike vendor software, which is specially tuned to specific detectors, BARNI must work well for a wide range of spectroscopic gamma-ray detectors. Concurrent with that goal is ease of use, which is necessary in order for community adoption of the tool as a benchmark for radionuclide identification. This document defines the software requirements needed to meet both these objectives.

2 Algorithm Overview

The BARNI algorithm has three broad objectives:

1) Perform radionuclide identification on a diverse set of instruments (e.q. NaI, HPGe).
2) Require minimal effort to setup and run analysis by the end-user.
3) Provide a benchmark for algorithms implemented in commercial systems.

To define what is meant by an "algorithm" in the context of radionuclide identification before proceeding to the strategies of alleviating the tension among the three objectives for BARNI software, consider the two main parts of radionuclide identification software which utilizes peak finding:

1) **Peak finder:** Find peaks and other features (e.g. Compton edge) in the spectrum.
2) **Expert system:** Use the information gathered by the peak finder to classify the measured radionuclides contributing to the spectrum.

Peak finding algorithms require detector characteristic information (e.g. energy resolution, efficiency), while expert systems require either manual or automatic training. The majority of expert system used in commercial RIIDs are manually trained (hard coded), and thus are both time consuming to construct and specific to one detector type. In order to meet the objective of BARNI software, an automatic training system will be engineered to create the expert system from detector representative samples with minimal user input. The diagram of the relationship between the algorithm, its configurations and training tools is shown in Figure 1.

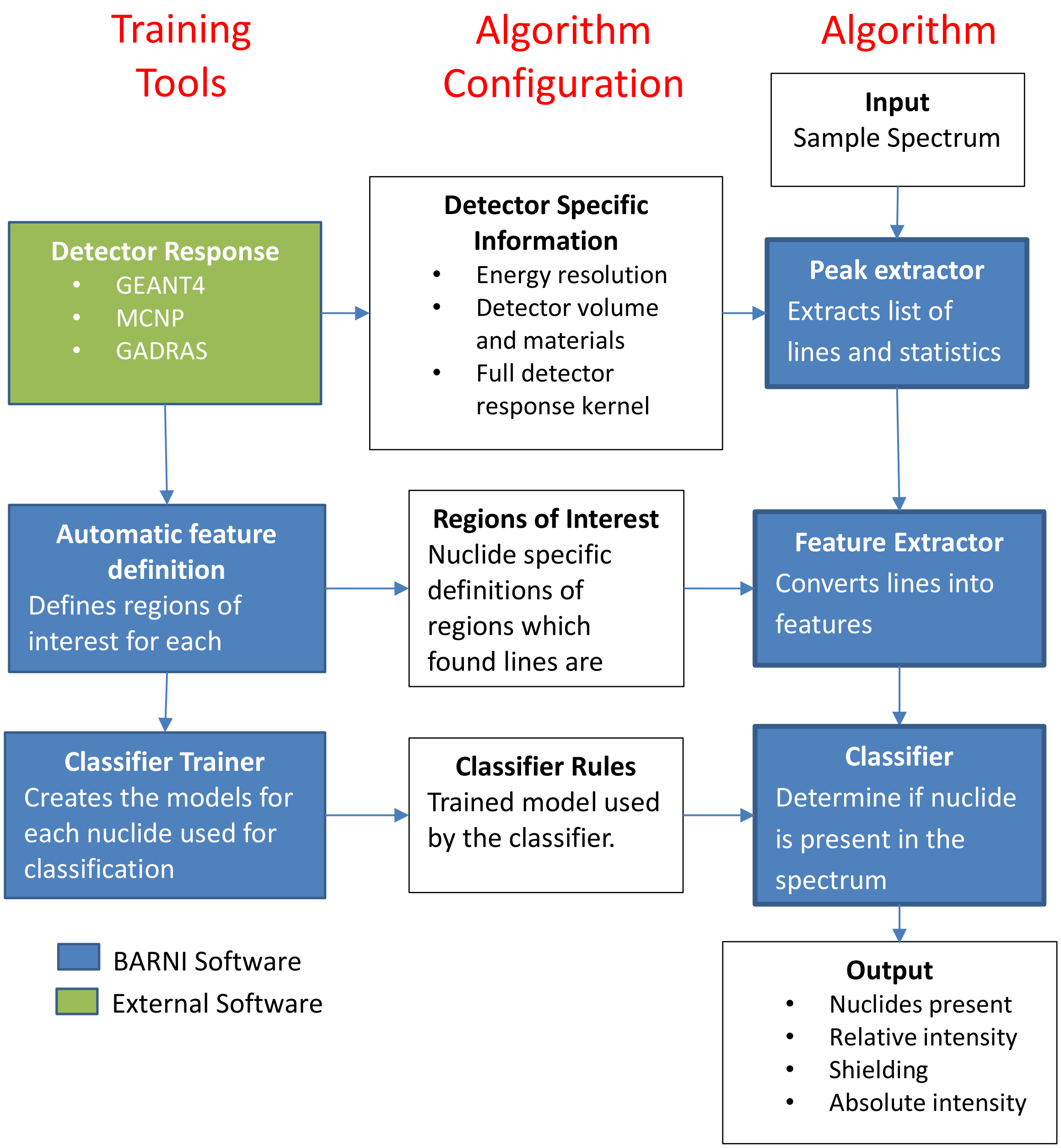

Figure 1 Overview of the BARNI algorithm and inputs. The colored boxes denote software, while the blank boxes indicate various files.

3 Product Scope

The scope of the BARNI project is to develop open-source radionuclide identification software for the analysis of gamma-ray spectra that provides the following functionality:

- Provide a list of likely radionuclides that contribute to a measured gamma-ray spectrum
- Integrate with RASE via the command line interface.
- Provide a benchmark for comparison with other radionuclide identification algorithms, such as commercial vendor algorithms for RIIDs, with sufficient common elements for a meaningful comparison.

The required user specified inputs include:

- Sample spectrum (in energy space), with live and real measurement time.
- All sample input formats accepted by RASE.
- Detector specific information, ranging from detector resolution to full detector response function.

The outputs of the automated expert system generation tools will be:

- Feature extraction rules which define the regions of interest into which located peaks are grouped into.
- Trained classifier model for each radionuclide of interest.

4 Product Functions

The software will automatically provide a list of the radionuclides present, and the relative intensity of each radionuclide, given an input spectrum and aforementioned required setup inputs.

4.1 User Characteristics

BARNI will be an open-source software, allowing for unrestricted use by anyone. Only basic computer skills will be required to use the algorithm for identification.

5 Architecture and Design

5.1 User Interface

- Command-line interface interoperable with RASE.

- Appropriate classes and functions shall be exposed as an Application Programing Interface (API) for integration of BARNI with other software.
- The core functionality of the BARNI algorithm shall be separate from the command-line interface.
- Specific widget toolkit for creating graphical user interface is to be determined, but it shall be free and open-source.
- Input requirements shall be provided as either human and machine-readable text formats (e.g. XML) or serialized objects (e.g. Python's pickle).

5.2 Programming Language

BARNI shall be developed and released in Python, which will leverage the languages strength in fast prototyping and a vast library of relevant open-source toolkits (e.q. scikit-learn). Any performance critical functions should be ported to compiled langue modules (e.g. C or Java).

5.3 Distribution

- The software should run from compiled executables but may require compilation by the end user.
- The software shall come with build system to facilitate compilation.
- The software shall run on a standalone desktop computer (e.g. quad-core 2GHz CPU with 4 GB of RAM) with reasonable execution times.
- The software shall ship with trained expert system for generic instruments (e.g. 3"x3" NaI)

5.4 Operating Environment

- The software shall install and run on any of the three major desktop operating systems: Windows, Mac OS, and Linux.
- The software's dependency libraries shall be freely available.
- The software shall not require separate installation of third-party software, outside of openly available Python libraries (e.g. numpy, scipy).

5.5 Design and Implementation Constraints

- The software shall have a command-line interface.
- The software shall interface with RASE software (see Appendix 2).

- The software documentation shall define abstract structures of the necessary parts of the algorithm (e.g. input, detector models, features, classifiers, results etc.).

6 Functional Requirements

6.1 Training

The expert system for a specific detector has to be built using a provided automated training system. The following are a list of requirements for this automated training system:

- The software shall include a set of supplemental tools for defining the input rules for both the feature extractor and classifier.
- The software shall include supplemental tools for running external detector response tools (e.g. Geant4, MCNP, GADRAS) necessary for generating training samples.
- The user should complete full training in a reasonable time frame (less than a week) with moderate CPU resources (e.g. commercial desktop workstation).
- The training samples should encompass a useful range of operation (e.g. observable counts, attenuation, mixtures, and background).

6.2 Operation

This section defines the functionality of the main radionuclide identification algorithm of BARNI.

- The software shall produce a list of nuclides present in the spectrum.
- The software should provide probabilities associated for each radionuclide found.
- The software output shall be able to identify up to 3 unique simultaneous radionuclides.
- The software shall give some metric to serve as quality of match in terms peaks covered.
- The software should associate each peak found with a radionuclide.
- The software shall be able to read a common input format ANSI N42.42 (2012 release) provided the file contains gamma spectrum measurement (RadMeasurement), detector specifications (RadDetectorInformation).
- The software shall be able to produce a common output format ANSI N42.42 AnalysisResults (2012 release) .
- The software shall use standard formats for input specifications (e.g. XML), or serialized objects with existing readers and writers.

6.3 Testing

This section defines the functionality of the set of helper tools for testing and scoring radionuclide identification performance.

- The software shall have test bench that exercises the entire training range of radionuclides and shielding configurations.
- The test bench shall produce a confusion matrix for a specific set of conditions (counts, shielding range, sources).
- The test bench shall produce an F-score from the confusion matrix.
- The test bench should produce a shielding range coverage plot showing the accuracy as a function of atomic number (Z) and areal density (AD) for a fixed count condition.

7 Additional Requirements

7.1 Licensing

- The software shall be release under an MIT license.
- The open release of the software shall comply with all Department of Energy and LLNL rules and regulations.
- The software shall comply with all export control regulations provided by the Department of Commerce (see Appendix 1)

7.2 Maintainability of Programming Code

- The software shall follow a common agreed upon style guide (e.g. PEP-8).
- The software should include mechanism for providing error and associated log messages in an event of a crash.
- The software should define abstract base classes for all core functions of the identification algorithm.
- The software shall be version controlled with well-known source control tool (e.g. git).
- The software major and minor releases should be marked with an appropriate version numbering scheme.

7.3 Documentation

- The software API documentation shall be autogenerated from in-code annotations.
- The software shall provide a build scripts for document generator (e.g. Sphinx).
- Developers shall provide a User Guide along with example problems.
- Developers shall provide Release Notes with each major version update.

7.4 Validation and Verification

The software shall undergo a formal verification and validation in order to ensure proper functioning of the radionuclide identification algorithm and supporting software. The verification shall ensure that:

- Single nuclide spectrum is properly processed.
- Anomalous spectrum input (zero live/real time) fails gracefully.
- Detector specifications are read in properly.
- Feature rules are read in properly.
- Classification rules are read in properly.
- Radionuclide identification output is correct and complete.
- Training tool output is correct and complete.
- Test bench output is correct and complete.
- Documentation is autogenerated properly.

Appendix 1: Export Control Guide

The proposed Benchmarking Algorithm for RadioNuclide Identification (BARNI) will be an open-source software tool designed for radionuclide identification in gamma-ray spectra. It will serve as a general use tool, and reference for the community for assessing the performance of identification algorithms. The following document outlines the export control guidance for radionuclide identification algorithms, and the requirements for BARNI to ensure open-source publication.

The Department of Commerce requires export license for any item described by a particular Export Control Classification Number (ECCN). According to the Bureau of Industry and Security (BIS) radiation detectors fall under the following ECCNs found on the Commerce Control List (CCL):

1) Hand-help first responder detectors (ECCN 1A004… with national security controls)
2) "General purpose" radiation detectors, such as Geiger counters, dosimeters (ECCN 1A999…. with only Anti-Terrorism (AT) controls)
3) "Everything in between" (ECCN 2A291… with nonproliferation controls, specifically designed for detection of SNM).

Software related to nuclear detection fall under ECCN 2D290, quoting from the Commerce Control List:

> *"Software" "specifically designed" or modified for the "development," "production," or "use" of items controlled by 2A290 or 2A291.*

The software is **NOT** deemed "specifically designed" if it is developed as a general purpose software (i.e. no "knowledge" of a particular commodity), or being developed with "knowledge" that it would be in use with commodities described in an ECCN controlled for AT-only reasons (i.e. 1A999) or EAR99 commodities. Any item that is not listed on the CCL is designated EAR99.

BARNI is designed to meet the above definition of "general purpose", since the published identification algorithm is designed to work for wide-range of detectors without detector specific inputs which have to be provided by the user, as shown Figure 1. Therefore, BARNI would be Export Controlled under EAR99, with only restrictions placed on embargoed/terrorist/sanctioned countries. With no Nuclear Nonproliferation restriction, the software can be made open-source provided DOE rules are followed.

However, even if we were to provide detector-specific radionuclide identification application, it will not fall under ECCN 2D290 if the specific instrument is not "specifically designed" for detecting or measuring special nuclear material. The relevant part of ECCN 2A291 covers detection of special nuclear material:

Radiation detectors and monitors "specifically designed" for detecting or measuring "special nuclear material" (as defined in 10 CFR Part 110) or for nuclear reactors.

As a result, even if the software were tuned to work with a particular detector, it would still not fall under ECCN 2D290 if the algorithm treats all radionuclides the same regardless of category. The following two examples of software that would fall outside of any ECCN and therefore only be Export Controlled under EAR99 rules:

1. Software designed and tuned for the FLIR R400, including detection of all commonly included radionuclides (including U235, U233, Pu239), but with no special designs for SNM specifically, would not fall under 2D290/2A291.
2. Software not designed or tuned for any specific instrument, but specially designed for the measuring SNM, would not fall under 2D290, and therefore not subject to 2A291.

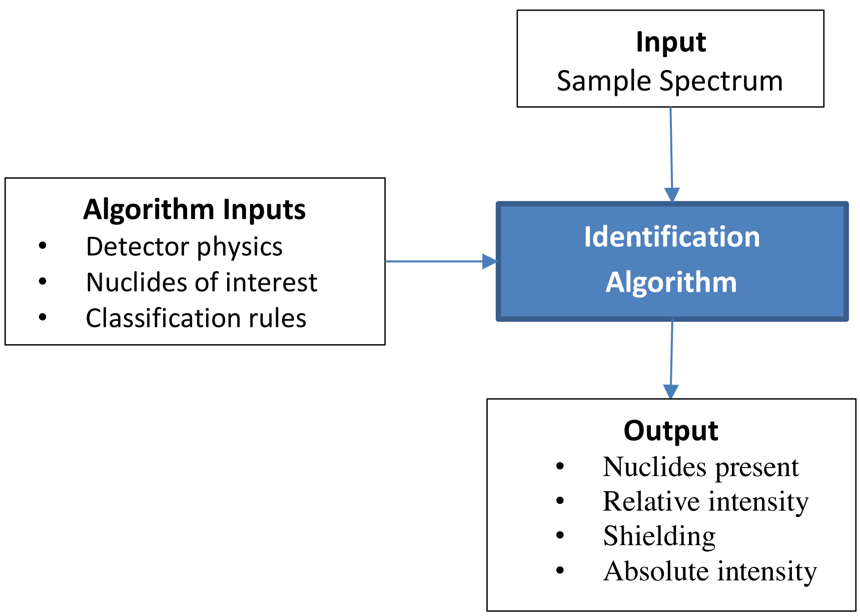

Figure 2 Simplified diagram of BARNI architecture.

Appendix 2: Identification algorithm requirements for RASE

- The identification algorithm shall be compatible with Windows 7 or above, or recent Mac OSX or Linux operating systems
- The identification algorithm shall allow for batch operation from the command line. At a minimum, it must accept two arguments, one for the path of the folder containing the spectra to process, and one for the path of the folder where identification results will be saved. Command line options shall follow POSIX specifications.
    - Example: utility_name INPUT_PATH OUTPUT_PATH
- The replay tool shall include a complete manual for installation, operation and troubleshooting.
- The replay tool shall include documentation for each parameter setting.
- If the algorithm allows for adjustable settings, a list of recommended (default) settings shall be provided.
- The replay tool documentation shall include the file input and output architecture description. This includes complete description of the data file formats and content. Example input and output files shall be also provided.
- The identification algorithm shall process spectra in standard ANSI N42.42 format
- The identification algorithm shall produce a single result file for each input spectrum.
- The file name for output file containing identification results shall be the same as an input sample spectrum with extension '.res'
- The data in the identification results files shall be ASCII characters. The data shall be encapsulated using the extensible mark-up language (XML).
- The file shall provide the list of isotopes identified and corresponding confidence index – a measure of how reliable the indication of the given isotope in the results of identification is, expressed as a number from 1 to 10 that indicates the degree of certainty to which result is correct.
- If statistics are too low to give a conclusive result, the indicator 'LOW' shall appear in the file. If radionuclide(s) is not recognized, the indicator 'UNKNOWN' shall be given in the identification results file.

- Example of the identification result file is given below:

```
<?xml version="1.0" encoding="UTF-8"?>
<IdentificationResults>
     <Identification>
          <IDName>K-40</IDName>
          <IDConfidence>9</IDConfidence>
     </Indentification>
     <Indentification>
          <IDName>Ra-226</IDName>
          <IDConfidence>4</IDConfidence>
```

</Indentification>
</IdentificationResults>