CTBTO
PREPARATORY COMMISSION

Peak identification in EDS Measurements using multiple subset sum problem formulation

Mohamed MAIZA

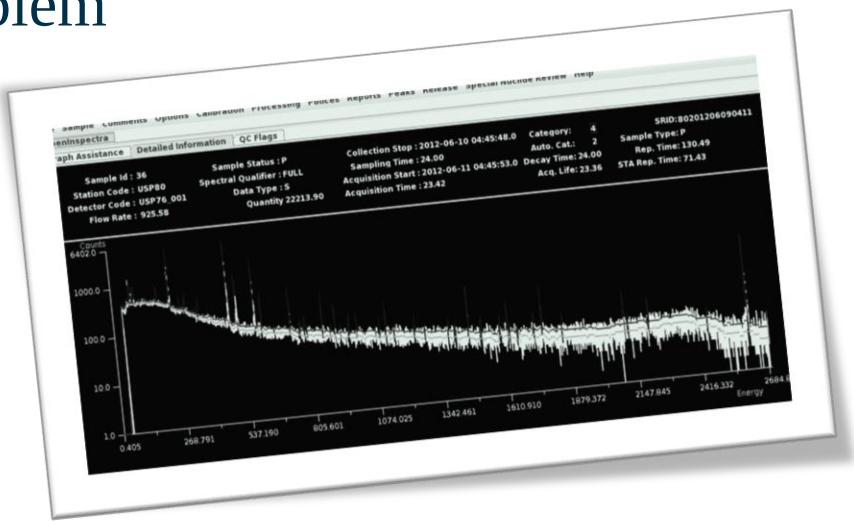
National Data Centre of Algeria

Modeling and Optimization Technics Laboratory
Military Polytechnic School of Algiers
BP 17 Bordj-El-Bahri, 16111 Algiers
ALGERIA.

CTBT: Science and Technology 2019 Conference
Hofburg Palace, Vienna, Austria
24 to 28 June 2019

Outline

- Context and Background
- Subset-Sum Problem - SSP
- Gamma Spectrometry
- Unassociated Peak Vs Radionuclide
- Solving the Peak identification Problem
- Summary



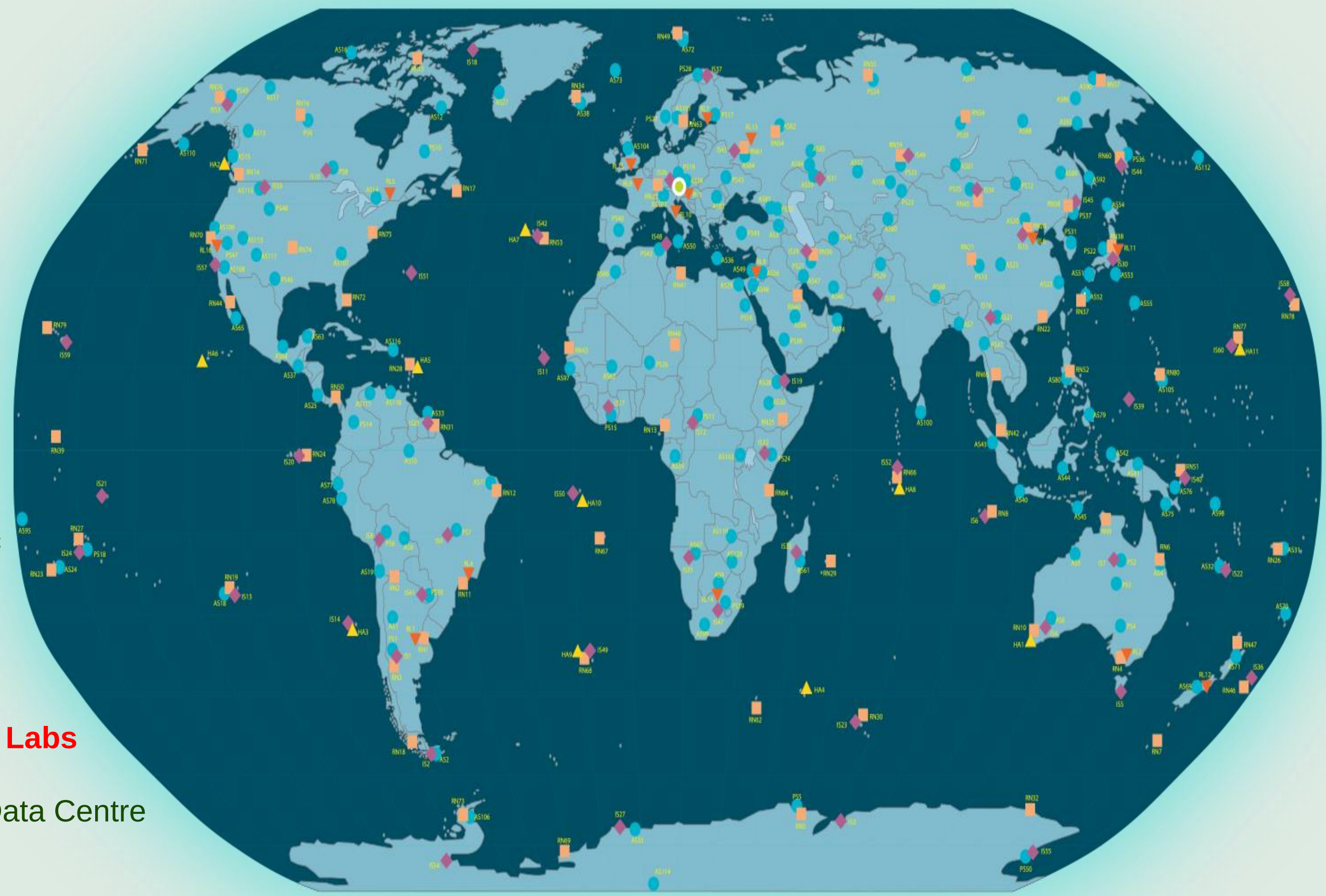
CTBTO PREPARATORY COMMISSION

- In the CTBT context, an "event" is any physical occurrence that is registered by the International Monitoring System, be it
  - ✓ *a natural or a man made event,*
  - ✓ *an earthquake or an explosion,*
  - ✓ *a chemical or a nuclear explosion.*

With regards to determining the nature of an event, Members States make the final judgement.

CTBTO PREPARATORY COMMISSION

International Monitoring System: 337 Facilities



- Seismic
- Hydroacoustic
- Infrasound
- **Radionuclide**
- **Radionuclide Labs**
- International Data Centre

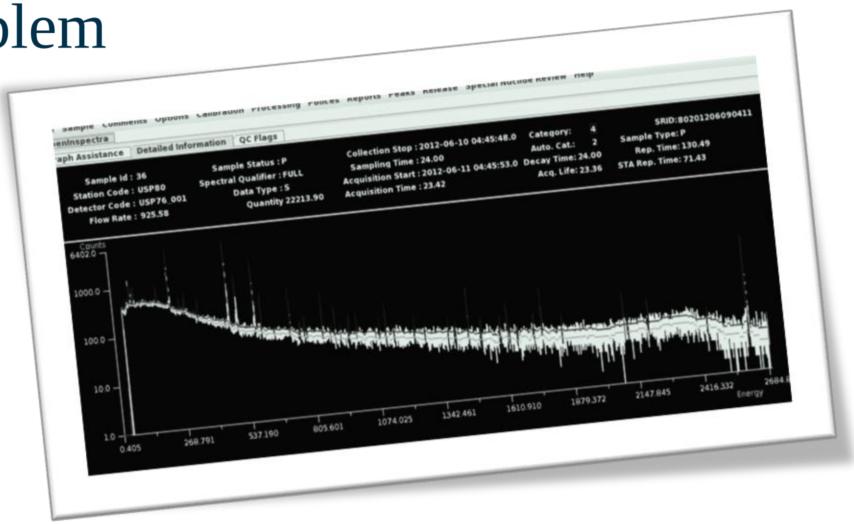
- Waveform data help identify the location of any event and qualify it as either natural or potentially man-made.
- However, they cannot reliably answer the question as to whether a man-made event, such as an explosion, was nuclear or not. The only reliable way to answer this question is to trace and **analyse** the radioactive remains of a potential nuclear explosion.
- Some remains provide the ultimate proof for the nuclear nature of an event
- Radionuclide monitoring data from the 80 radionuclide monitoring stations worldwide are expected to shed light on this key question.
- Data on radionuclide observations are sent to the International Data Centre (IDC), where they undergo an analysis process.
- After the automatic analysis process, analysts refine the results during interactive review.

- Only the analysis of radionuclide data can provide the ultimate proof to establish the nuclear nature of an explosion
- Each radionuclide particulate monitoring station sends one gamma ray spectrum per day showing which radionuclide were observed in a single sample and in what quantity.
- Each radionuclide Sample was examined and categorized in the corresponding group according to the nature of the radionuclide contained in them.

- In the spectrum analysis process, doubts about the existence of level 4 or 5 nuclides are made.
- In order to ensure that a **suspect nuclide** exists or not, this presentation addresses the problem of **association of unidentified peaks with suspect nuclides**.
- We assume that the corresponding energy is an amount of **sum of energies** provided by the **suspects nuclides**.

Outline

- Context and Background
- **Subset-Sum Problem - SSP**
- Gamma Spectrometry
- Unassociated Peak Vs Radionuclide
- Solving the Peak identification Problem
- Summary



- The Subset Sum Problem (SSP) is an important problem in computer science and combinatorial optimization.
- The challenge is to determine if there is some subset of numbers in an array that can sum up to some number ***s***.
- SSP is to find subset of elements that are selected from a given set ***S*** whose sum adds up to a given number ***K***, where we consider that the set ***S*** contains non-negative values.

Given :

$\boldsymbol{A = (a_1, a_2, \ldots, a_n)}$ and

a positive integer **s**, called the sum.

The question is then : Is there a subset ***A'*** of ***A*** with

$\boldsymbol{SUM_{a' \text{ in } A'}(a') = s}$, or equivalent:

Does there exist a vector

$\mathbf{X = (x_1, x_2, \ldots, x_n), x_i \text{ in } \{0, 1\}, \text{ with } AX = s?}$

If a solution is desired, the problem becomes a **functional problem**:

The functional problem of any vector ***A*** together with a sum ***s*** is denoted by **(A, s).**

CTBTO PREPARATORY COMMISSION

Input Data

- A = (15, 22, 14, 26, 32, 9, 16, 8)
- s = 53.

Solution

(A, 53) is given by the vector X = (1, 1, 0, 0, 0, 0, 1, 0), because

$AX = a_1 + a_2 + a_7$

$= 15 + 22 + 16 = 53.$

The corresponding subset is A' = (15, 22, 16).

it's not the only solution

- Generally, there may exist several solutions to (A, s)
- Another solution for (A, 53) is X = (0, 1, 1, 0, 0, 1, 0, 1) with the subset A' = (22, 14, 9, 8).

- The SSP is in the NP complexity class.
- The decision problem is NP-complete and the corresponding functional problem is NP-hard.
- No polynomial algorithm is known solving the general subset sum problem.
- The trivial possible solution vector for a given (A,s) is :

for all possible 0-1 vectors $X = (x_1, ..., x_n)$ the sum $s' = AX$ is computed.

But, this algorithm is computing the sum of all ***2^n*** different ***X*** vectors, and has an exponential running time of ***$O(2^n)$.***

Meet-in-the-middle algorithm

- One of the more efficient algorithms;
- Recursive algorithm;
- At each iteration, this algorithm divides the original input vector A in two equal parts $A_1 = (a_1, ..., a_t)$ and $A_2 = (a_{t+1}, ..., a_n)$.
- All possible sums $s_1 = A_1X_1$ are computed in a pre-computation step.

CTBTO PREPARATORY COMMISSION

Meet-in-the-middle algorithm

- Input: a vector $A = (a_1, ..., a_n)$ and a sum s.
- Output: a vector X with $AX = s$, provided a solution exists.

I- Divide $A = (a_1, ..., a_n)$ into $A_1 = (a_1, ..., a_t)$ and $A_2 = (a_{t+1}, ..., a_n)$.

For every vector $X_1 = (x_1, ..., x_t)$

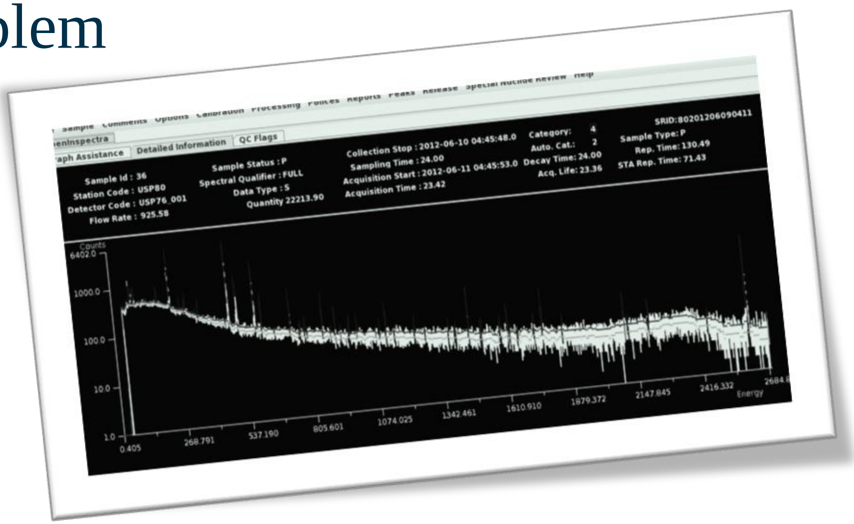
- ➢ compute $s_1 = A_1X_1$
- ➢ store s_1 together with the corresponding ***x-vector*** in a table ***T***
- ➢ sort this table by the sum

II- Now for all possible vectors $X_2 = (x_{t+1}, ..., x_n)$

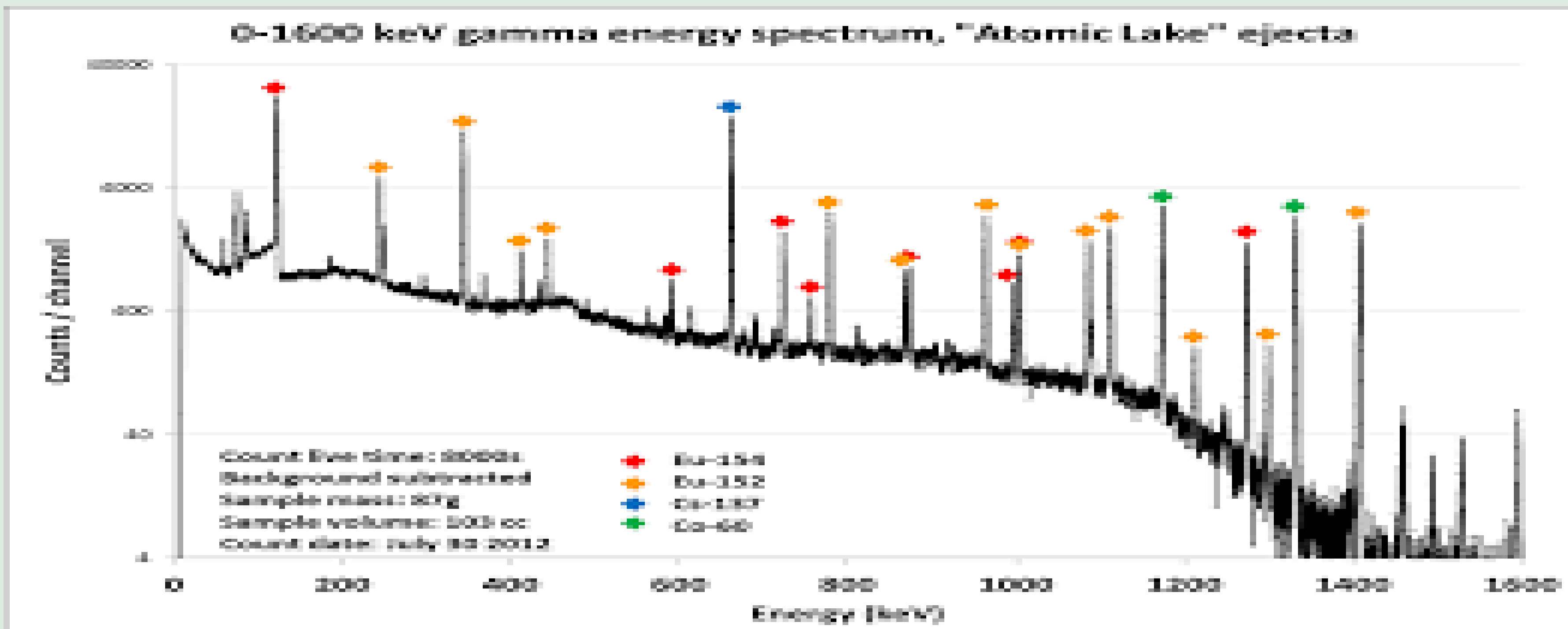
- ➢ compute $s_2 = A_2X_2$
- ➢ get the difference $L = s - s_2$
- ➢ search L in T. If found, a solution vector is $X = X_1||X_2$.

Outline

- Context and Background
- Subset-Sum Problem - SSP
- **Gamma Spectrometry**
- Unassociated Peak Vs Radionuclide
- Solving the Peak identification Problem
- Summary

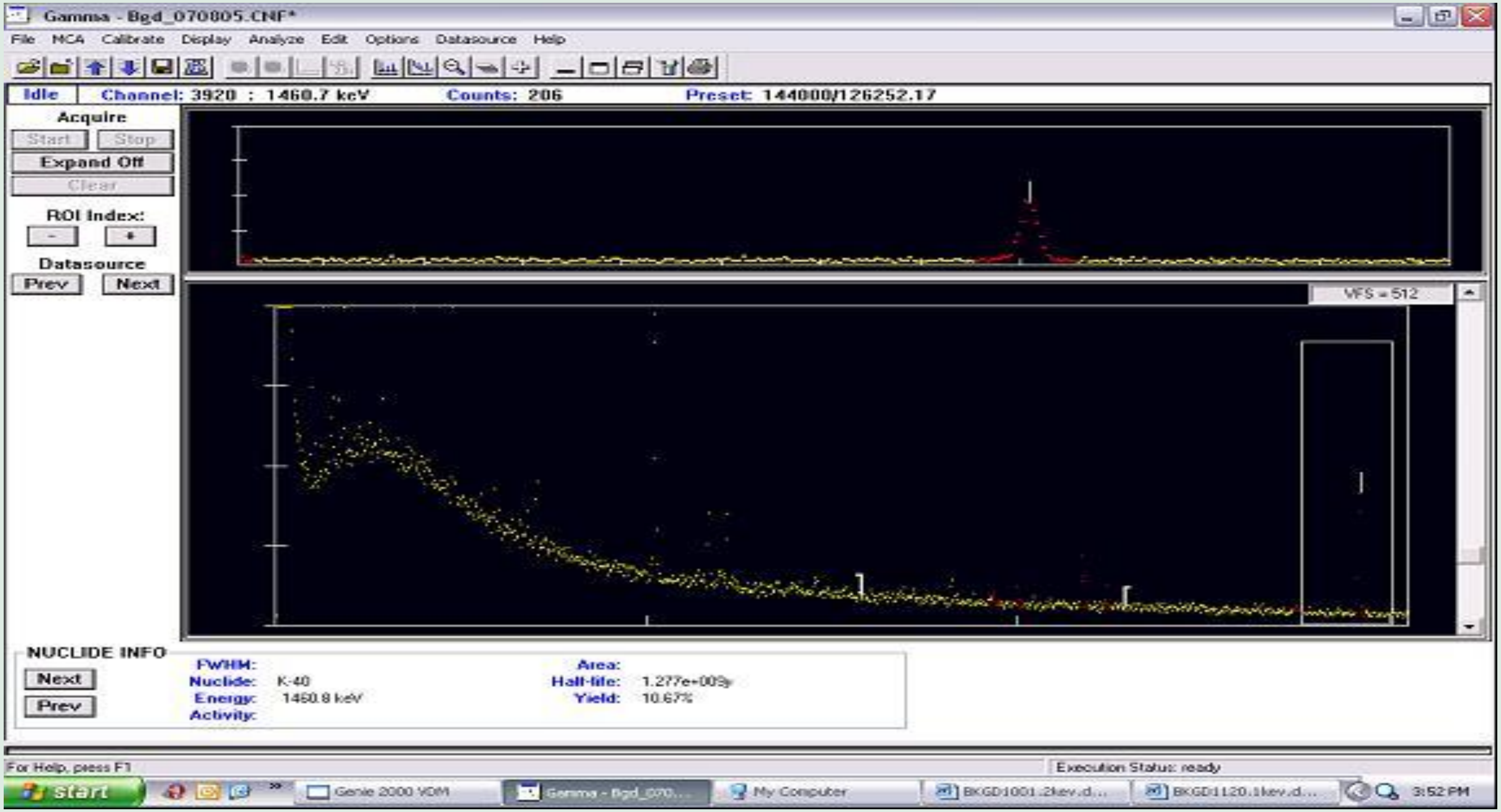


- Most radioactive sources produce gamma rays, which are of various energies and intensities. When these emissions are detected and analyzed with a spectroscopy system, a gamma-ray energy spectrum can be produced.



- Gamma spectrometry is one of the fundamental measuring techniques in nuclear technology for unambiguous, clear identification of radionuclide, even in mixtures of various isotopes.
- The experiment is intended to show the basics of the methodology of gamma spectrometry, the properties of most important gamma detectors and how to identify clearly radionuclides.
- Moreover, the absolute activity of a radionuclide could be determined by means of a gamma spectrometer if it is calibrated absolutely with regards to its efficiency.

- For an IDC or NDCs Analyst, the Analysis of the identity of gamma-emitting radionuclides and measurement of their respective activity in a sample are the important tasks



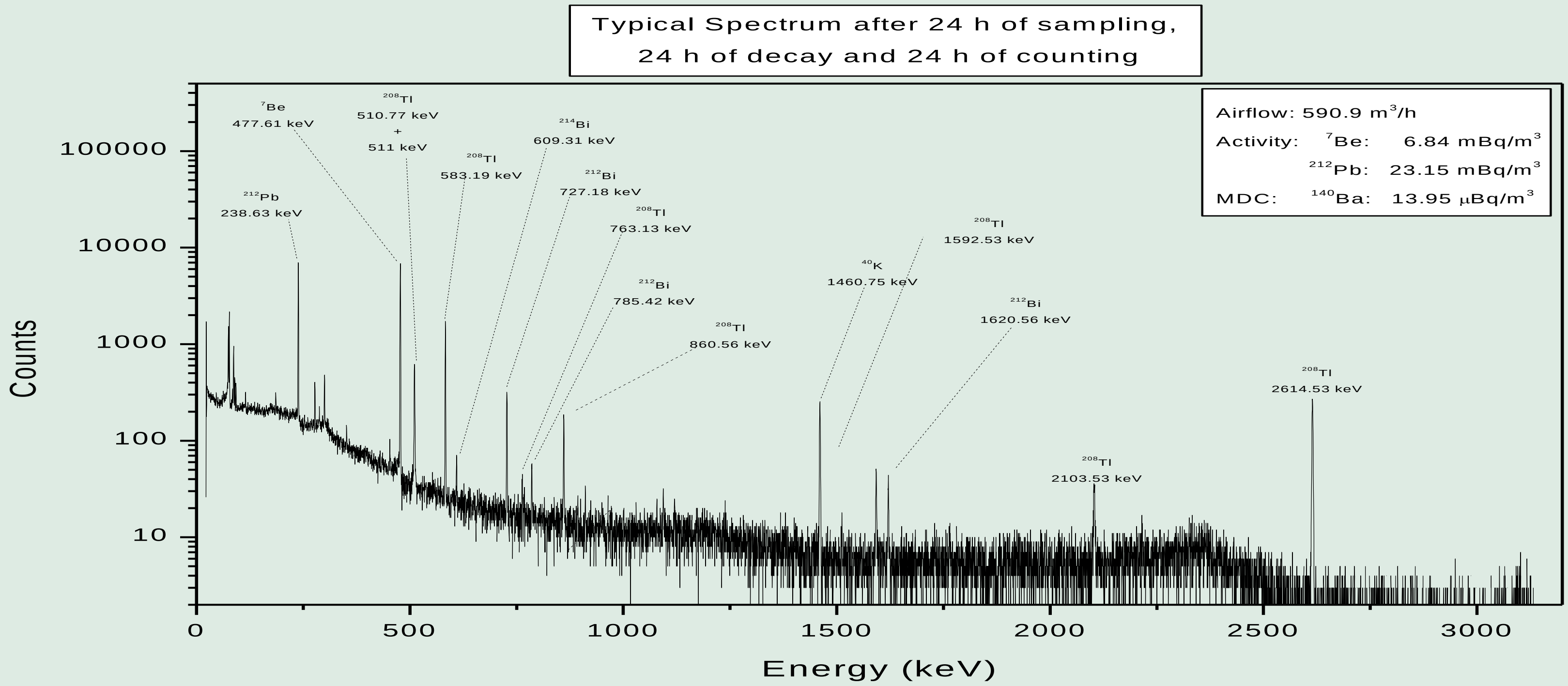
Peak identification using SSP formulation

CTBTO PREPARATORY COMMISSION

Energy

Gamma emitters can be identified by their respective energies

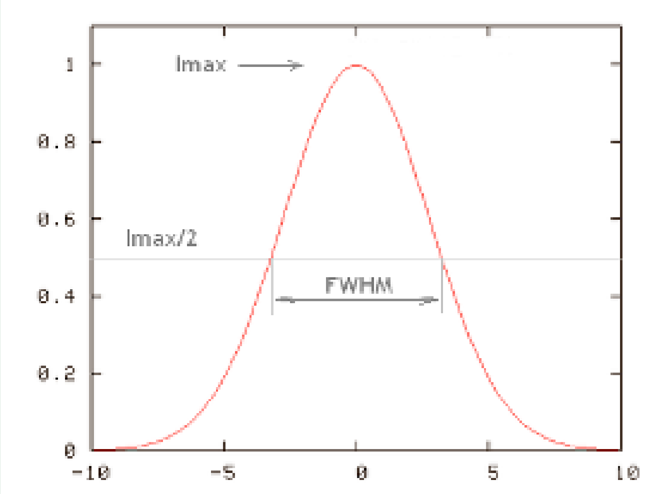
Important characteristic for detection – very low energy gammas are difficult to detect; gamma emitters with similar energies are difficult to identify



Detectors characteristics

Resolution

- ability to separate two gamma peaks that are close to each other in order to be able to separately measure and identify these peaks.
- usually measured with the full-width at half maximum (FWHM). This is the width of the gamma ray peak at half of the highest point on the peak distribution.
- A requirement for gamma detectors used in radionuclide stations is that the FWHM should be <2.5 keV at 1332 keV (Co-60 peak)



Efficiency

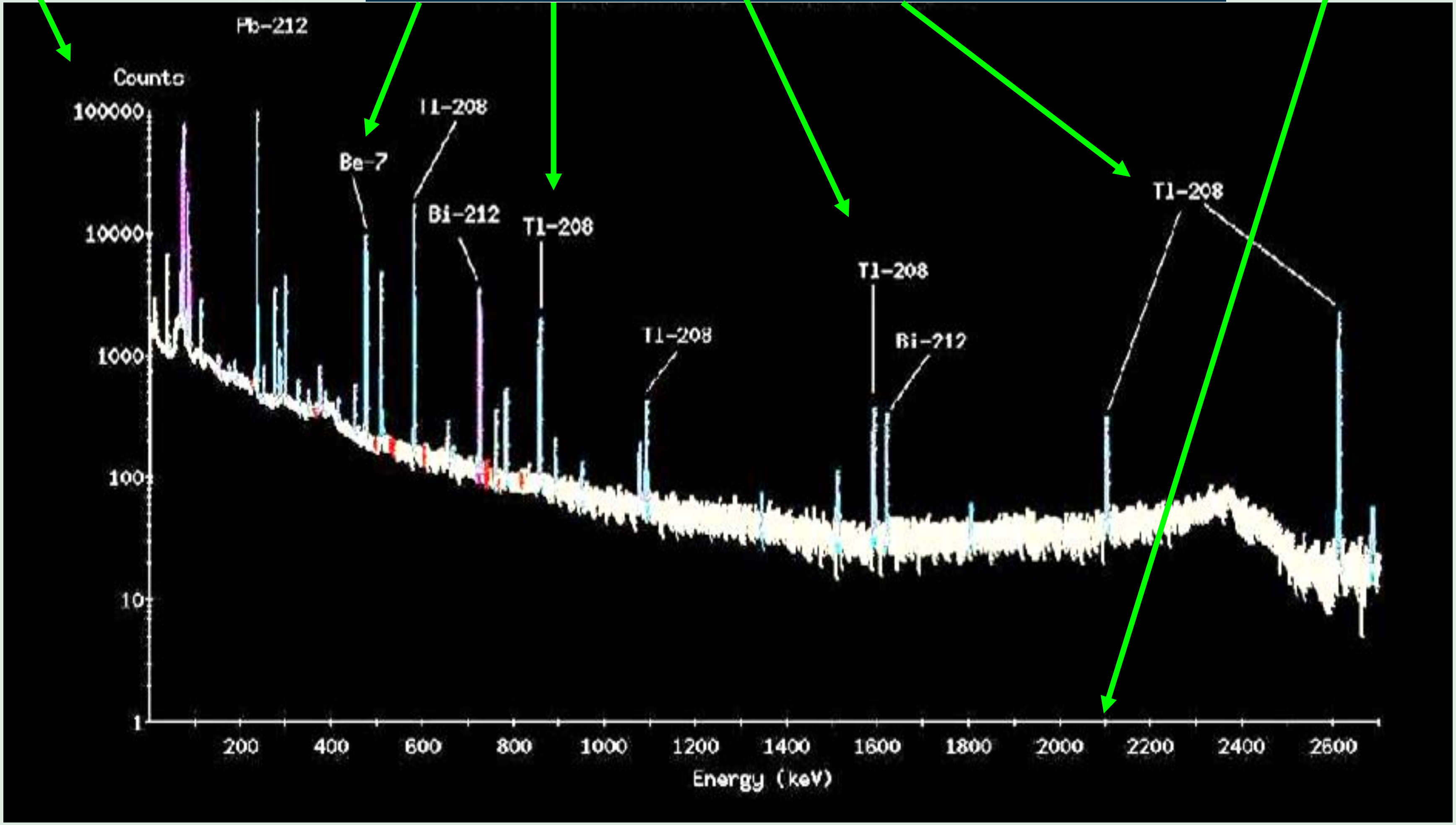
- Relative efficiency values are often used for germanium detectors, and compare the efficiency of the detector at 1332 MeV to that of a 3 in × 3 in NaI detector (another gamma detector). This type of detector was popular and well characterized before the introduction of the germanium detectors.
- Relative efficiency of detectors used at radionuclide stations should be ≥ 40%. The higher the efficiency, the better the detection sensitivity.

Example of station sample spectra

Number of γ-photons collected in the detector

What energy do they have?

What radionuclides and how much are in the sample?



CTBTO PREPARATORY COMMISSION

Software for interactive review : OpenSpectra

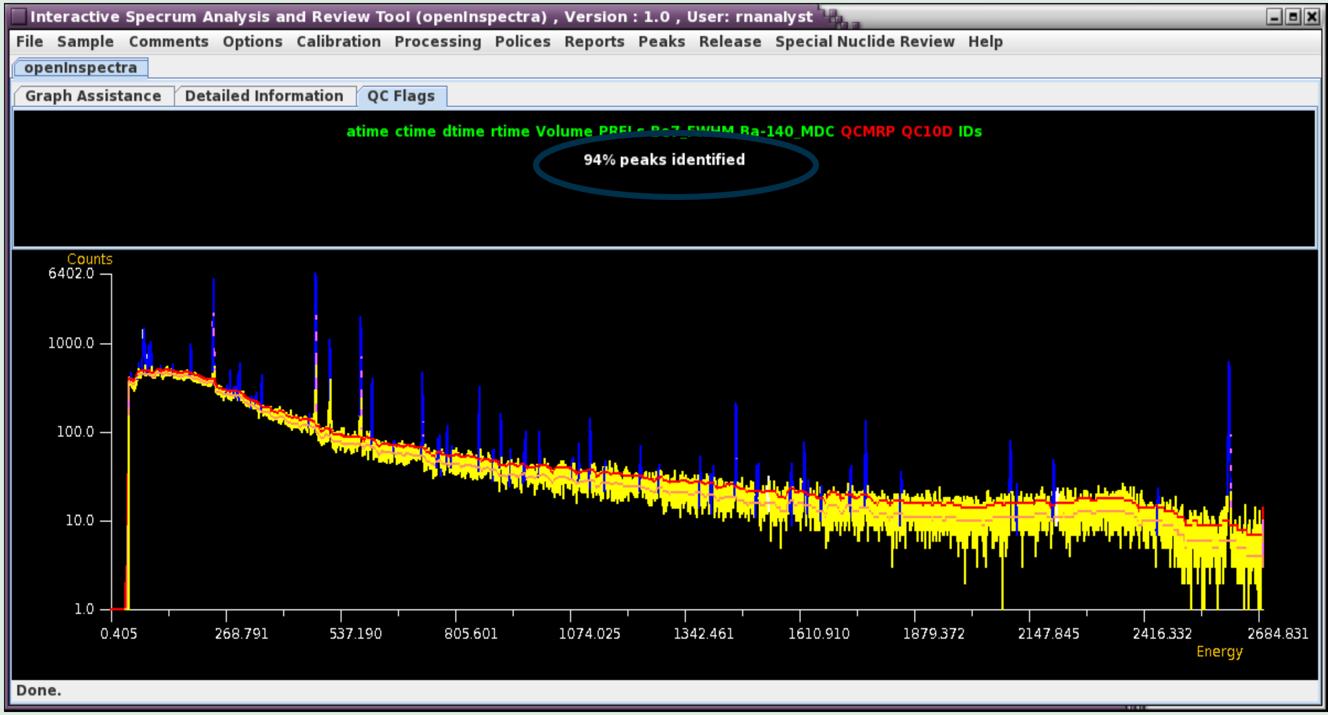
- A java-based license-free application for interactive review of automatic processing results which covers both particulates stations and Noble gas systems.

Sample metrics:

- Collection time
- Decay time
- Acquisition time
- Reporting time
- Be-7 FWHM
- Ba-140 MDC
- MRP QC
- QC 10 days
- %Identified peaks

Color coding:

- GREEN: *Pass*
- RED: *Fail*

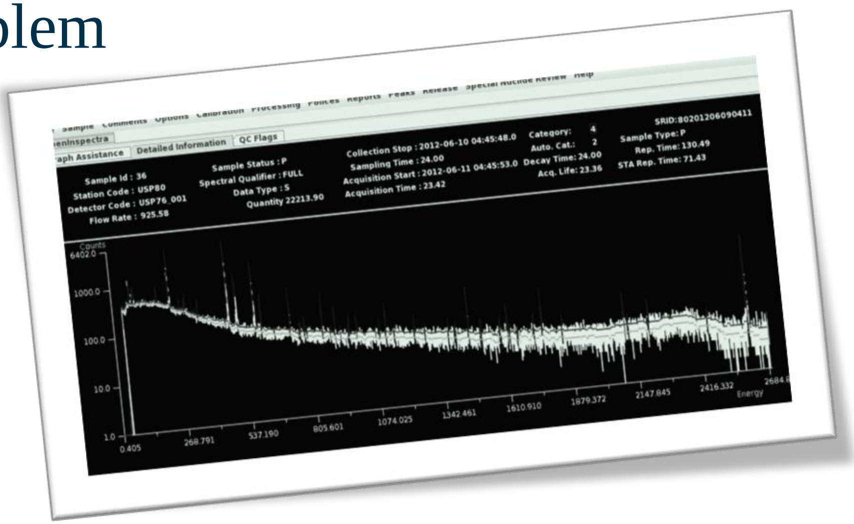


Software for interactive review

- Some of peaks (value of energy) are sometimes not identified;
- One of main analyst tasks is to affect the correct nuclide to the corresponding value of energy. This task allows the precise identification and quantification of the elements;

Outline

- Context and Background
- Subset-Sum Problem - SSP
- Gamma Spectrometry
- **Unassociated Peak-Radionuclide**
- Solving the Peak identification Problem
- Summary



CTBTO PREPARATORY COMMISSION

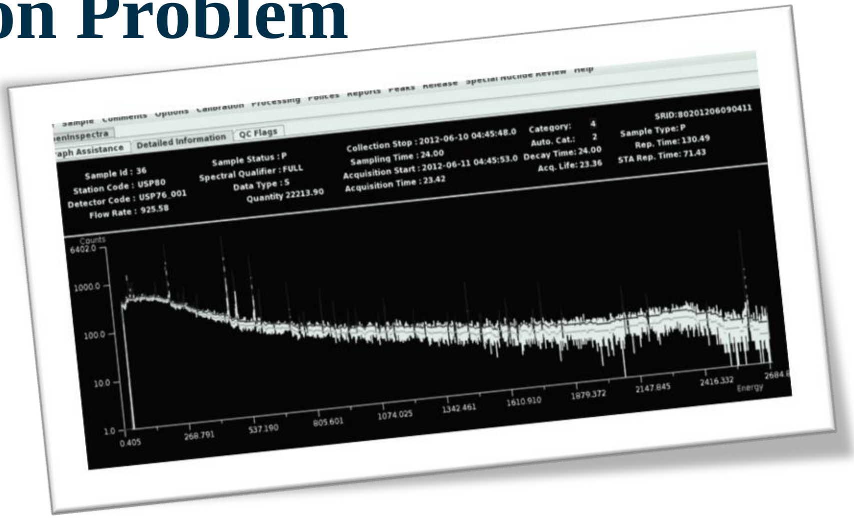
- One of main analyst tasks is to affect the correct nuclide to the corresponding value of energy. This task allows the precise identification and quantification of the elements;
- Some of peaks which corresponding to a given value of energy are sometimes not identified;
- The non-identification is due to the fact that the corresponding energy does not correspond to any classified nuclide;
- In order to identify such a peak and associates it with one or more radionuclide, we apply a formulation of a subset sum problem.

- In fact, there are some peaks that are already associated to the corresponding radionuclide (automatic identification peak).
- These radionuclides along with **suspect radionuclides** are candidates to be the source of energy corresponding to the unidentified peak
- The problem can be expressed by the following :

From already detected radionuclides and the set of suspect radionuclides, what are the radionuclides whose energy accumulation corresponds to the energy of the unidentified peak.

Outline

- Context and Background
- Subset-Sum Problem - SSP
- Gamma Spectrometry
- Unassociated Peak-Radionuclide
- **Solving the Peak identification Problem**
- Summary



CTBTO PREPARATORY COMMISSION

- This problem can be formulated as a **functional problem** ***(A,s)*** of a **subset sum problem** as follows:

Input :

A : the vector which contain energy value of all existing radionuclides, together with energy value of suspect radionuclides

s : The energy of the unidentified peak (the peak which is not yet associated with a radionuclides)

Output :

The set of energy value selected (radionuclide detected).Then, the peak become identified

CTBTO PREPARATORY COMMISSION

- Let $A = (a_1, \ldots, a_n)$ a set of energy value of all existing radionuclides, together with energy value of suspect radionuclides
- Let $X = (X_1, \ldots, X_n)$ a set of boolean decision variable where

  $X_i = 1$ if the corresponding energy selected

  $X_i = 0$ *otherwise*
- Given s the energy of the unidentified peak

Identification of the peak consist on selecting a subset of X whose total corresponding energies is closest to, without exceeding, the value s

- Integer linear program ILP formulation

$$\text{Maximize } z = \sum a_i X_i$$

Subject to

$$\sum a_i X_i \leq s$$

$$X_i = 0 \text{ or } 1$$

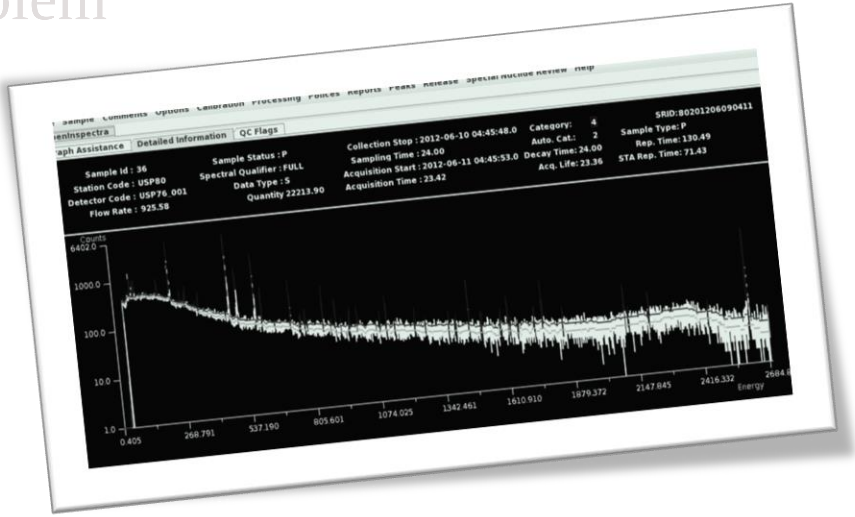
Many approach resolution to the SSP problem

- Exact Algorithms *Faaland (1973); Ahrens and Finke(1975); Martello and Toth (1984)*
  - Dynamic programming,
  - Branch and bound method
- Approximate Algorithms
  - Greedy algorithms,
  - Polynomial time approximation schemes
  - fully polynomial time approximation schemes
  - Probabilistic analysis
- Hybrid Algorithms

Outline

CTBTO
PREPARATORY COMMISSION

- Context and Background
- Subset-Sum Problem - SSP
- Gamma Spectrometry
- Unassociated Peak-Radionuclide
- Solving the Peak identification Problem
- **Summary**



- SSP formulation for the problem of assignment of radionuclides to any peak without accordance energy-radionuclide;
- Our idea allows the analyst to reduce the number of unidentified peaks
- The SSP can be solved in polynomial or pseudo polynomial time for small instances
- The SSP formulation is very effective for testing the existence of such radionuclide which have several levels of energy

Thank you for your attention