Disarmament Verification

Areg Danagoulian

MIT
New START treaty, 2011 – Russia & USA

• Reduce deployed warheads to 1550 warheads each -- ~3x reduction

• How do treaty partners verify that the other side is dismantling actual warheads and not fakes? They don’t.

• Verification: delivery vehicles – easier to verify.

• Problems: large leftover of non-deployed warheads
  • theft → nuclear terrorism, nuclear proliferation
  • rapid rearmament in times of political crisis

→ Authenticate warheads, without revealing classified information!
Overall View of Thrust Area

• Treaty verification is not the same as weapon detection.

• The goal of verification is to confirm that an object presented as “X” is “X”.
  – Negotiate protocols to establish acceptable level for “confirmation.”

• Critical Issues:
  – clear all real warheads (completeness)
  – detect all fakes/hoaxes (soundness)
  – reveal no classified information (“zero knowledge”)
Thrust Area V Subprograms

- Verification Using Inherently Trustworthy Instruments (Univ. of Michigan)
  - SAR ADC’s with non-uniform bin resolution
  - Lead: David Wehe
  - Student: Fred Buhler
  - Collaborating with: LLNL

- Information Barriers with Enhanced Automated Isotope Identification (UIUC)
  - Lead: Clair Sullivan
  - Student: Mara Watson,
  - Collaborating with: DAF

- Zero Knowledge template verification (Princeton + Yale)
  - neutron radiography \(\rightarrow\) comparison to a template
  - Leads: Alex Glaser, Francesco d’Errico, Robert Goldston.
  - Student: Sebastian Philippe
  - Collaborating with: PPPL, DAF

- Physical Cryptographic Verification of Nuclear Warheads (MIT)
  - transmission NRF to produce a physical hash of a nuclear warhead \(\rightarrow\) comparison to a template
  - Leads: Richard Lanza, Areg Danagoulian, Scott Kemp
  - Students: Jayson Vavrek, Ruaridh Macdonald, Ellie Immerman, Bobby Nelson, Jake Hecla
  - Collaborating with: PNNL
Verification Using Inherently Trustworthy Instruments

David Wehe
University of Michigan
For treaty verification, both parties must agree on a measurement protocol that provides adequate assurance that treaty obligations are met without yielding sensitive information.

Existing approaches use templates or information barriers applied post-measurement, and are suspicious because sensitive information is acquired before the barrier.

E.g.: FPGAs hackable, power changes detectable by untrusted observer.

This work investigates electronic measurement techniques in which precise, spoof-proof, digital information can be acquired only where mutually acceptable.
Successive Approximation ADC. Fast, high resolution.

- Pulse height is measured to a resolution of $2^{-n}$ after $n^{th}$ step.
- $n+1$ step only taken if could fit into a predefined range
- High resolution in ROI, no/low information away from allowable ranges.
- Significant gain in throughput
- No measurement of irrelevant information
New two-stage SAR ADC architecture

- First stage only detects if input is within an agreed upon spectrum
- Error amplifier limited to desired window
- Second stage digitizes the error signal

Untrusted Observer Immunity: Each stage is physically (capacitor size) and electrically (saturation) limited to the agreed upon spectrum

Side Attack Immunity: SAR decision tree removes power supply correlation with ADC code
Thrust Area V subprograms

- Verification Using Inherently Trustworthy Instruments (Univ. of Michigan)
  - SAR ADC’s with non-uniform bin resolution
  - Lead: David Wehe
  - Student: Fred Buhler
  - Collaborating with: LLNL

- Information Barriers with Enhanced Automated Isotope Identification (UIUC)
  - Lead: Clair Sullivan
  - Student: Mara Watson,
  - Collaborating with: DAF

- Zero Knowledge template verification (Princeton + Yale)
  - neutron radiography $\rightarrow$ comparison to a template
  - Leads: Alex Glaser, Francesco d’Errico, Robert Goldston.
  - Student: Sebastian Philippe
  - Collaborating with: PPPL, DAF

- Physical Cryptographic Verification of Nuclear Warheads (MIT)
  - transmission NRF to produce a physical hash of a nuclear warhead $\rightarrow$ comparison to a template
  - Leads: Richard Lanza, Areg Danagoulian, Scott Kemp
  - Students: Jayson Vavrek, Ruaridh Macdonald, Ellie Immerman, Amelia Turner, Jake Hecla
  - Collaborating with: PNNL
Information Barriers with Enhanced Automated Isotope Identification (UIUC)

• Given allowable peaks to be measured:
  • Enhanced automated isotope identification algorithms for improved information security
  • Accurate identification with different detectors
• Results:

![Graph showing energy spectrum](image)

Identifications with > 50% posterior probability:
Pu-239: 84.7%
Am-241: 84.1%
Pu-239 + Am-241: 57.7%

Spectrum of BeRP ball + 4 cm polyethylene + 1.27 cm Pb, collected with 2x2 in. NaI in 2 minutes
Wavelet Analysis and Derivation of Peak Areas

Step 1: Calculate wavelet transform

\[ WX = S \]

\[ S(a, E) = \int_{-\infty}^{\infty} X(t) \cdot \frac{1}{\sqrt{a}} \psi^*(\frac{t - E}{a}) dt \]

\( S \): = CWT coefficient matrix
\( W \): = Wavelet transform tensor
\( X \): = Spectrum vector
\( E \): = Wavelet centroid parameter
\( a \): = Wavelet scale parameter
\( B \): = Wavelet basis matrix
\( B_1 \): = Optimal submatrix of \( B \)
\( k \): = Fit vector, wavelet representation of \( X \)
\( C_S \): = Signal covariance
\( C_k \): = Fit covariance
Peak centroids, areas, uncertainties provided to Bayesian ID code

- Sample identifications made from DAF and LANL measurements
- All spectra collected with NaI, 60 second integration time
- Red indicates incorrect identification

<table>
<thead>
<tr>
<th>Source</th>
<th>Shielding</th>
<th>Amount</th>
<th>Distance (cm)</th>
<th>ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>Se-75</td>
<td></td>
<td>0.54 mCi</td>
<td>50</td>
<td>Se-75, Eu-152</td>
</tr>
<tr>
<td>Eu152</td>
<td></td>
<td>10 uCi</td>
<td>30</td>
<td>Eu-152</td>
</tr>
<tr>
<td>Eu152 / Ba-133</td>
<td></td>
<td>10 uCi /</td>
<td>10 / 2</td>
<td>Eu-152</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.4 uCi</td>
<td></td>
<td></td>
</tr>
<tr>
<td>U-233</td>
<td></td>
<td>1 g</td>
<td>100</td>
<td>U233, Th-232</td>
</tr>
<tr>
<td>HEU (93.2% U-235)</td>
<td>1.2 cm Fe</td>
<td>13 kg</td>
<td>120</td>
<td>U-235</td>
</tr>
<tr>
<td>WGPu (BeRP Ball)</td>
<td></td>
<td>4.5 kg</td>
<td>120</td>
<td>Pu-239, Am-241, I-125</td>
</tr>
<tr>
<td>(93.7% Pu-239)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WGPu (BeRP Ball)</td>
<td>1.2 cm Pb</td>
<td>4.5 kg</td>
<td>120</td>
<td>Pu-239, Am-241</td>
</tr>
<tr>
<td>(93.7% Pu-239)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Thrust Area V subprograms

• Verification Using Inherently Trustworthy Instruments (Univ. of Michigan)
  – SAR ADC’s with non-uniform bin resolution
  – Lead: David Wehe
  – Student: Fred Buhler
  – Collaborating with: LLNL

• Information Barriers with Enhanced Automated Isotope Identification (UIUC)
  – Lead: Clair Sullivan
  – Student: Mara Watson,
  – Collaborating with: DAF

• Zero Knowledge template verification (Princeton + Yale)
  – neutron radiography → comparison to a template
  – Leads: Alex Glaser, Francesco d’Errico, Robert Goldston.
  – Student: Sebastian Philippe
  – Collaborating with: PPPL, DAF

• Physical Cryptographic Verification of Nuclear Warheads (MIT)
  – transmission NRF to produce a physical hash of a nuclear warhead → comparison to a template
  – Leads: Richard Lanza, Areg Danagoulian, Scott Kemp
  – Students: Jayson Vavrek, Ruaridh Macdonald, Ellie Immerman, Amelia Turner, Jake Hecla
  – Collaborating with: PNNL
Zero Knowledge Warhead Verification with Neutron Transmission and Emission Measurements


- Use active neutron interrogation in a Zero Knowledge configuration:
  - transmission radiographs are recorded on detectors preloaded with the complement radiograph (including Poisson noise) of a reference item.
  - If the item is valid (identical to the reference), the final radiograph is identical to the expected exposure if no object had been present.

- Proof-of-concept system demonstrates fast neutron differential radiography can confirm that two objects have identical neutron opacity without revealing geometries/composition.

ZERO-KNOWLEDGE WARHEAD VERIFICATION

HIGHLIGHTS OF EXPERIMENTS

14-MEV OBJECT-COMPARISON SYSTEM @ PPPL

- 1st demonstration of a physical zero-knowledge proof
- Different configurations of 2’’ metal cubes
- Results show that when objects are identical, inspectors do not learn geometry or composition

ACTIVE INTERROGATION OF HEU @ DAF

- Transmission and emission measurements with different types of bubble detectors
- Two configurations of the Rocky Flats HEU shells
- Different sources with ~300-keV (AmLi), 2.5-MeV (DD) and 14-MeV (DT) neutrons
ZERO-KNOWLEDGE WARHEAD VERIFICATION
DETECTOR DEVELOPMENT (YALE)

**TRANSMISSION**
- Capable of storing > 1,000 counts
  - Preloads indistinguishable from measurement counts
- Insensitive to gamma radiation
- Sensitive to neutrons above selected thresholds
  - Some thresholds of interest: 3 and 10 MeV

**EMISSION** (spontaneous and driven)
- Capable of storing thousands of counts
  - No imaging at present
- Insensitive to gamma radiation
- Sensitive mainly to fission neutrons
  - Energy threshold ~500 keV or above source
Consortium for Verification Technology

ZERO-KNOWLEDGE WARHEAD VERIFICATION

NEW READING TECHNIQUES (YALE + PRINCETON)

OPTOELECTRONIC READOUT (YALE)
- A beam of infra-red light crosses the active area of the detector and is deflected by evaporated bubbles.
- Photodiodes affixed along the detector length selectively detect the scattered light component post-irradiation.

360-OPTICAL TOMOGRAPH (PU)
- Takes 360-degrees movies of detectors.
- Use PU open-source bubble counting software. (in development).
- To be upgraded with HeNe laser scattering for data commitment experiments.
OPEN SOURCE INFORMATION BARRIER
PASSIVE GAMMA AND NEUTRON

INFORMATION BARRIER EXPERIMENTAL (IBX)
- Open source software, towards open hardware
- Encourage others to improve or defeat IBX
- Successfully tested at DAF

MULTI-CRITERIA-TEMPLATE APPROACH
- Compares gamma spectrum and count rate
- Compares neutrons indirectly through 2223 keV gammas from polyethylene → sensible to mass
- Implemented in IBX

Thrust Area V subprograms

• Verification Using Inherently Trustworthy Instruments (Univ. of Michigan)
  – SAR ADC’s with non-uniform bin resolution
  – Lead: David Wehe
  – Student: Fred Buhler
  – Collaborating with: LLNL

• Information Barriers with Enhanced Automated Isotope Identification (UIUC)
  – Lead: Clair Sullivan
  – Student: Mara Watson,
  – Collaborating with: DAF

• Zero Knowledge template verification (Princeton + Yale)
  – neutron radiography \(\rightarrow\) comparison to a template
  – Leads: Alex Glaser, Francesco d’Errico, Robert Goldston.
  – Student: Sebastian Philippe
  – Collaborating with: PPPL, DAF

• Physical Cryptographic Verification of Nuclear Warheads (MIT)
  – transmission NRF to produce a physical hash of a nuclear warhead \(\rightarrow\) comparison to a template
  – Leads: Areg Danagoulian, Scott Kemp, Richard Lanza
  – Postdoc: Brian Henderson
  – Students: Jayson Vavrek, Ruaridh Macdonald, Bobby Nelson, Jake Hecla
  – Collaborating with: PNNL
**NRF Weapon authentication Concept**

**Weapon B: candidate**

- Bremsstrahlung
- NRF-filtered brem
- Everything classified by the host
- Everything open

**Shielding**

- tNRF signal: “hash”

**Physical Cryptography:**
- No direct data from the weapon itself
- \[\text{SIGNAL} = (\text{Weapon}) \otimes (\text{Foil})\]
- Impossible to extract (weapon)

**Soundness and completeness:**
- Authenticated template A -- acquire \(S_{NRF}(A)\)
- Candidate weapon B -- acquire \(S_{NRF}(B)\) and compare
can catch most hoaxes within minutes

Verification Concept with transmission NRF

- Preliminary results:
  - $^{238}$U – observed all the primary and secondary resonances
  - Al – acquired data for normalization to $^{27}$Al’s known cross section

- Students and postdocs
  - Jayson Vavrek
  - Ruaridh Macdonald
  - Dr. Brian Henderson (Stanton postdoctoral fellow)

- Collaboration with Ken Jarman, PNNL, on the information theory problem
General framework for comparing warhead verification protocols

- Quantify how each step of the protocol effects completeness/soundness/secrecy
- Methods developed from techniques in problems of data privacy

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_{c</td>
<td>u}$</td>
</tr>
<tr>
<td>$p_{x,s</td>
<td>c}$</td>
</tr>
<tr>
<td>$p_{y</td>
<td>x}$</td>
</tr>
<tr>
<td>$f(Y)$</td>
<td>Function describing the verification protocol</td>
</tr>
</tbody>
</table>

Unmeasurable user information

- Latent properties of the test object
- Measurable data about the test object
- Secret & unused data

Private, measurable test object information

- Data which will be used in the verification protocol

Released, measurable test object information

- Transformed / scrambled test object data
- Pass / fail indicator variable

Axiom linking latent properties (e.g. ‘warheadness’) and measurable data

Function which extracts the data which host and investigator agree is necessary to identify a warhead via attributes or template

Function describing the process by which the released data is obfuscated to protect warhead secrets

Function describing the verification protocol
Conclusion

- Solid progress on all projects:
  - experimental proof of concept demonstration of neutron radiography verification protocol (Princeton)
  - optimized neutron bubble detectors for Zero Knowledge neutron radiography (Yale)
  - completed feasibility simulations of the Nuclear Resonance Fluorescence (NRF) protocol, taking experimental data (MIT)
  - analyzed data from Device Assembly Facility (DAF) for spectral algorithm development (Illinois)
  - Developing a new, non-uniform ADC concept for gamma spectroscopy (UM)
BACKUP
Understanding the problem

Fundamental Analog to Digital Converters (ADCs)

<table>
<thead>
<tr>
<th>Resolution [Bits]</th>
<th>Bandwidth [Hz]</th>
</tr>
</thead>
<tbody>
<tr>
<td>ΣΔ</td>
<td>Configurable</td>
</tr>
<tr>
<td>SAR</td>
<td>Pipeline</td>
</tr>
<tr>
<td>Flash</td>
<td></td>
</tr>
</tbody>
</table>

Consortium for Verification Technology
Moving Forward

- Exploits existing R&D for consumer appliances.
- State of the art, published in ISSCC 2015
- SAR ADC with 50 Msps, 11.5 ENOB, 1mW, in CMOS
- 2-stage ADCs are common approach.
- Newly proposed architecture implementable with modifications to residue amp and 1\textsuperscript{st} SAR capacitor DAC
- Modify the current design to fabricate and test a candidate treaty-acceptable inspection system
- HPGe measurement system assembled.
- Interfaces with national lab, industrial partners during design phase.
We are developing a general framework for comparing warhead verification approaches

- Protocol steps are described by mathematical functions
  - All one-to-one / many-to-one relationships are made explicit
- Various measures of mutual information are used to quantify how much information about the test object / warhead is passed on at each protocol step
  - This quantifies the information the inspector receives, i.e. the completeness and soundness
  - The warhead owner can calculate the protocol secrecy using the mutual information between the measured data and the warhead design → this is agnostic to inference method
NUMERICAL LIMITS ON WARHEADS WITH “BUDDY TAG”

WARHEAD COUNTING

- The challenge: establish a baseline count of warheads and enforce a ban on un-tagged items in a variety of operational environments.
- Numerical counts of items must be trustable and the information security concerns of inspected parties must be respected.

BUDDY TAG CONCEPT

- Buddy tag acts as a companion token, proving ownership of a treaty-accountable item while remaining physically detached from the item itself.
- Declarations are verified by short notice inspections which confirm that all items are associated with a companion tag.
Flash ADC as example

- design with variable width comparators by adjusting the resistive ladders at the chip level.
- do this in CMOS technology to be cost effective.
- Key idea is to get superb energy resolution in regions of interest while blocking design information from scrutiny.
- Snag: Available flash ADCs do not have sufficient bit resolution. Need $10^3$-$10^4$ matched comparators.