Leveraging the Wisdom of the Crowd: Hardware and Software Challenges for Nuclear Disarmament Verification

Moritz Kütt, Sébastien Philippe and Alexander Glaser

July 12, 2015

Abstract

The verification of nuclear disarmament will require the use of adequate measurement technologies. All will need to enable inspecting parties to trust the measurement chain leading to valid and correct results, while ensuring host countries that classified information–especially warhead-design information–remains secure. However, this seems particularly difficult with technologies that rely heavily on electronics and software such as most radiation measurement system. Crowdsourcing and open-source challenges are a potential source of innovation to address this issue. This paper presents results from verification challenges experiments that were conducted in a classroom environment. Tasks included the development of simple information barriers and possible exploits for measurement electronics. From the results, the paper highlights requirements and propose ideas to scale-up those challenges from the classroom to larger audiences.

1 Background

Deep reductions in the global nuclear arsenals may be facilitated with verifiable agreements among the nuclear weapon states. Other actors such as non-nuclear weapon states and international organizations may wish to join such agreements as monitoring parties.

Nuclear weapons disarmament verification may require new technologies, including reliable measurement techniques, trusted and non-intrusive systems for data processing and protection and, efficient inspection protocols designed for this purpose. If some of the major concepts have been proposed in the 1990's and early 2000's, this research area is becoming very active again. One result of this recent

renewed interest has been the development of collaborations between nuclear-weapon states laboratories and non-nuclear weapon states organizations as well as open and transparent partnerships with universities and non-governmental organizations. This new synergies have already produced major successes. We think that expanding the pool of participants even further could benefits innovation.

In this paper, we point out new ways of researching and testing innovative ideas for disarmament verification with an emphasis on so-called "crowdsourcing" and public innovation challenges. We present results from small scale verification challenges experiments that were conducted in a classroom environment at Technische Universität Darmstadt (Germany) and Princeton University (USA), and discuss ideas to scale-up those challenges from the classroom to larger audiences.

2 Public Innovation Challenges

The idea of public challenges is rather old, for example the "Longitude Rewards" established through an act of the British parliament in the 18th century encouraged the development of more accurate methods to determine ships longitude. The basic concept behind these challenges aims to leverage knowledge and creativity of many [6]. Over recent years, many companies as well as government authorities and non-governmental organizations have proposed challenges, with financial incentives or recognition prizes, and obtain significant results. Different definitions for crowdsourcing exist (cf. [2]). In general, challenges differ by time frame, prize, task, participation and result sharing. It is often important that the pool of participants reaches a significant size and diversity. Larger incentives are often required to address complex problems that would require years of work (e.g. million dollar clay institute milenials prize in mathematics). Smaller tasks can be addressed in specific venue such as "hackathons", where people meet and work on projects only for a couple of days or even hours, and present immediate results. Topics address a variety of problems including both technical and societal issues. We list here some relevant examples:

- Robotic Vehicles: DARPA Grand Challenge (`http://www.darpa.mil/NewsEvents/Releases/2014/03/13.aspx`)
- Electronic Voting: Voting Competition 2007 (`http://www.wired.com/2007/07/uncle-sam-wants/`)
- Solar Cell deployment: SunShot Challenge (15 Mio. $ prize, `http://energy.gov/eere/sunshot/funding-opportunity-announcement-solar-market-pathways`)
- Government Independent Space Flight: Ansari X Prize (`http://space.xprize.org/ansari-x-prize`)
- Hacking iCorruption, Edmond J. Safra Center for Ethics at Harvard University and the MIT Civic Media Center (`http://ethics.harvard.edu/hacking-icorruption`)

An additional form of challenge ask participants to find "bugs" and vulnerabilities in hardware and software, a very useful tool when developing technical solutions to address sensitive problems such as disarmament verification.

Particular for disarmament verification, challenges could center around a basic problem: handling of verification activities that involve information imbalances. Such imbalances are for example information that is only known to a host state, but can not be shared with an inspector state because of classification issues or treaty regulations. Often, information barriers are proposed as a solution. They would process sensitive information and output only non-sensitive data, sometimes limited to a single bit. To ensure that inspector and host would both trust these machines, a joint development process might be necessary. Several information barriers have been developed in the past, mainly by U.S. national laboratories and during the UK-Norway Initiative (cf. [1], [7], [3]).

In the following section, we present two cases were we took up the problem of information barrier development with participants that had no previous knowledge on the issue.

3 Verification Challenge Experiments

3.1 Princeton Verification Challenge

In Spring 2015, we offered a new course at Princeton University entitled *"Unmaking the Bomb: The Science and Technology of Nuclear Nonproliferation, Disarmament, and Verification"* (MAE 354/574), open to both undergraduate and graduate students. As part of this course, we proposed a six-week verification challenge that involved the development and construction of a number of prototype information barriers using the template-matching approach for use with a radiation detection inspection system based. Students were divided into three teams (two undergraduate teams; one graduate team).

Figure 1: Development kit (left) and a finished information barrier (right).

Each team started with the same development kit (Figure 1, left) and a small budget for extra hardware and material. All systems were based on a *Raspberry Pi* single-board computer connected a sodium-iodide detector unit. At the outset of the exercise, each team received a "secret" template file representing the gamma signature of their own nuclear "warhead" (Figure 2). In the development process, the teams had to make sure that design choices for their inspection systems would enable reliable recognition of their templates without leaking any of its features. In addition to the technical work, the teams were tasked to jointly negotiate in several in-class sessions the procedures for the inspection exercise and determine all essential features of their inspection systems.

Each project was graded based on the functionality and performance of the information barrier, the simplicity of the concept and, in particular, the ability of the other parties to authenticate the hardware and software. Extra credits were awarded to teams that were able to learn details about other templates or managed to "cheat" during an inspection without being caught. Unknown to the others, the graduate-student team was required to cheat during the mock inspections to receive full credit.

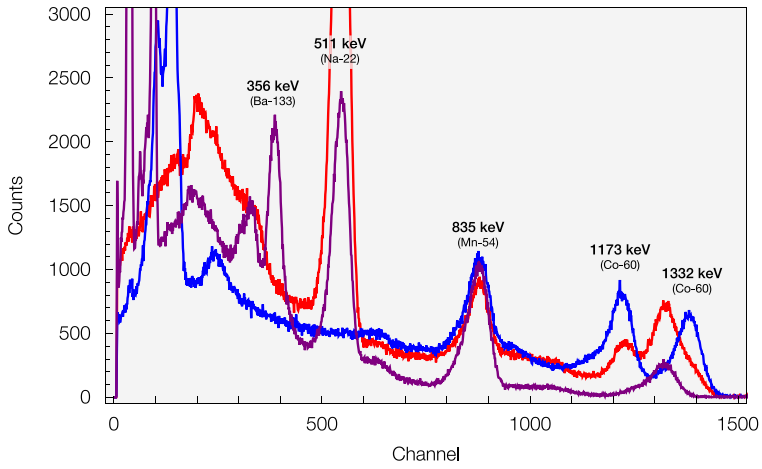

Figure 2: Template spectra for the Princeton Verification Challenge.

Verification Protocol. We anticipated that the students would follow "conventional wisdom" and adopt an approach where the host supplies the information barrier to be used for inspection of its own warheads. Especially, the team that had to cheat was counting on this choice as it was planning to implement a "hidden switch" to secretly control the box during inspections. During the negotiations, however, the other teams became increasingly suspicious about the intentions of the third team as it was requesting unusual procedures. Eventually, after much discussion of advantages and disadvantages of different approaches,

a new majority view emerged to have *inspector-supplied* information barriers instead. The host would be allowed to thoroughly inspect a box in the presence of the inspectors prior to use, the main hardware (i.e., the *Raspberry Pi*) would be acquired through blind buy, the software would be open source, and an agreed procedure would be used to demonstrate that the correct software is running during the inspection. Finally, only the host would operate the inspector-supplied box following the instructions provided by the inspector.

This radical departure from typical approaches, where the host de-facto controls the inspection equipment, essentially turns the common "host-advantage" into an "inspector-advantage." While the inspector is now confident that the system is working as intended, the remaining key efforts are now aimed at ensuring the host that no data is leaked during the inspection.

Technology. Essentially, the information barriers had to compare two low-resolution gamma-spectra and autonomously decide if both are equivalent, i.e., produced by the same radiation source. Algorithms running on the information barrier could essentially make no assumptions about the expected spectral features as each team only knew its own template spectrum. Beyond that, teams were only allowed to work with two calibration sources (cesium-137 and zinc-65) to test their hardware and software during development.

Eventually, all three teams converged on two basic design choices. First, templates were not stored on secure external storage media as proposed for other systems; instead, each template was generated anew prior to inspection and temporarily stored within the information barrier. This was considered acceptable since the boxes were intended for "one-time use" only and would remain with the host after the inspection. Second, all teams employed basic statistical tests to analyze and compare the radiation spectra from the template and the inspected item. In the hands-on lab sessions, the teams independently confirmed that the standard Kolmogorov-Smirnov test distinguished different calibration sources with high confidence and within very short time periods. This particular test measures the difference between two empirical cumulative distribution functions (CDFs) using, in this case, the radiation spectra as probability density distributions (counts versus channel). Eventually, all three systems performed successfully in several rounds of mock inspections in the last week of the semester.

3.2 Darmstadt

A course with the title "Messelektronik Hacken" ("Hacking Measurement electronics") was held at Technische Universität Darmstadt. About 15 students from differenct faculties (mainly physics, but also engineering and computer science) participated.

Scenario. For the seminar, a fictious scenario was invented: The governments of two countries, Rekah and Urilmar, jointly negotiated a disarmament agreement ("Dragondell agreement"). The agreement allowed Urilmar to confirm the authenticity of 30 nuclear weapons from Rekah's stockpiles. It is assumed that

Urilmar has prior confidence that one of these items is a nuclear weapon ("Golden Sample"). Both parties agreed to use the template approach to match remaining items with the Golden Sample using neutron counting mechanisms. For each item, two neutron counters will be placed at fixed positions (A, B). Counting rates of equal positions for Golden Sample and a Test Item will be compared by an Information Barrier, which will display a green light if counting rates of both items are equal at position A and B respectively.

Technology. The measurement setup was realized using a system based on the CAMAC standard (Computer Aided Measurement And Control). A four channel scaler (QUAD SCALER SR1608 from GEC-Elliott Automation Unlimited) was used with input from the four dectector positions. The scaler was placed in a CAMAC crate and was read out using a CAMAC crate controller and a set of two standard computer. CAMAC devices have been used for more than forty years, and the scalers and crate used in this exercise had a similar age. This hardware was chosen for two reasons: a) a large stock of devices was available in the nuclear physics department (a "trusted third party") and b) they were based on 70ies electronics with easily accesible printed circuit boards with through-hole mounted parts and integrated circuits in dual in-line packageformat (DIP/DIL) only.The Information Barrier was realized by a software program that ran on one of the computers. It was assumed that this computer was trusted by all parties.

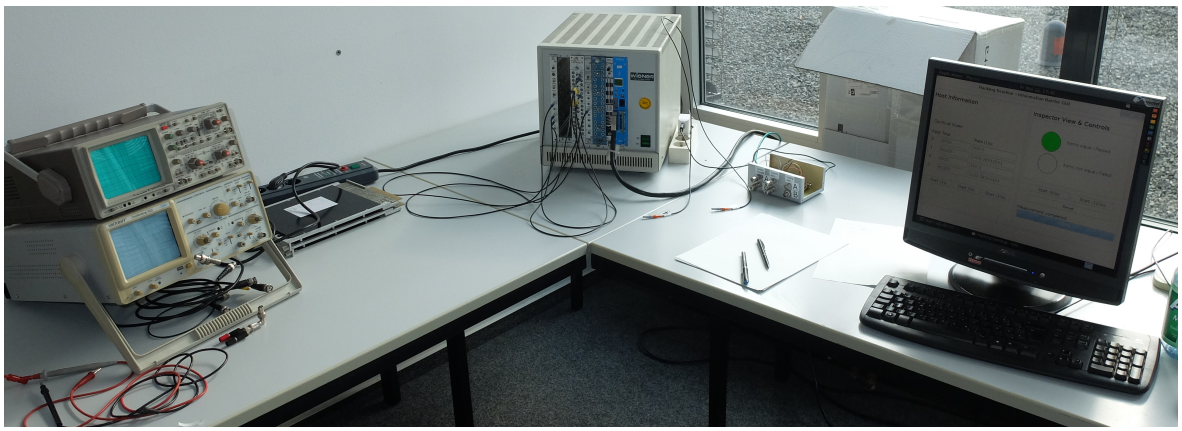

Figure 3: Laboratory setup in Darmstadt. The CAMAC crate is in the middle, on the right the computer displaying the software-based information barrier and between the "MeasureSim" device to simulate detector signals.

Because it was not possible to use radioactive sources in the rooms available, simulated signals were used instead of actual radioactive samples and detectors (cf. figure 3). A device was specifically developed for the course ("MeasureSim") to produce detector output pulses. A Correct Item or the Golden Sample would result in neutron count rates of 300 pulses/s at detector position A and 450 pulses/s at detector position B.

The students played the roles of scientists and engineers from the host state (Rekah) and were tasked to manipulate the Scaler. Their goals was to enable invalid items ("failing items") to successfully pass the inspection. These items had count rates for positions A and B of 400 pulses/s and 800 pulses/s respectively. The modification should not influence other measurements, and no parts of the setup other than the Scaler could be modified.

The students had four weeks to work on their ideas (parallel to others lessons) and were divided into 3 teams. Every team was handed out a working Scaler device in the beginning, students always had access to the full setup to test their modifications. Each team had a small budget (100 Euros) to buy parts needed.

The described setup might be simpler than what may be used for real inspections. However, it should be realistic enough to yield a broad range of modifications, that will need to be addressed in the future.

The students were instructed to keep the modifications as small as possible, however none were invisible. It was assumed that miniturization could have been easily realized with a larger budget. All groups used programmable microcontrollers (e.g. Arduino) for their solutions. Two of three groups managed to sucessfully manipulate the scaler. One group added a controller which counted pulses parallel to the actual scaler. When count rates similar to a failing item were detected, the counter was reset prior to readout and a fictious count rate was set. The second group rebuild a CAMAC bus interface which replaced the interface of the original scaler. Every read and write command was transmitted via the new interface, which allowed for easy manipulation whenever particular count rates were read.

At the end of the course, all groups presented their results. The students feedback was very positive. They enjoyed learning about measurement technologies in a context much different from a typical physics classroom. Motivation was high over the whole period. Positive influence on the motivation came probably from the hands-on experience of working with a real scaler device.

3.3 Lessons Learned for Verification

The "crowd" that participated in both exercise was small and problems were accessible to knowledge and capacity of undergraduate and graduate students. However, they came up with innovative ideas to improve or manipulate disarmament verification devices.

The most surprising outcome of the Princeton exercise, involving both host and inspector parties, was the choice of inspector-supplied equipment that would remain in host custody afer inspection (one-time use). This choice challenges conventional wisdom and turns the authentication/certification challenge upside down. Essentially, it solves one problem (convincing the inspector that the system is working) at the expense of exacerbating another (convincing the host that the system is not leaking information). Our experience suggests that the latter problem might be easier to resolve [5].

In both exercises, students were challenged by the complexity of the setup. It is important to have clearly documented specifications and limitation of the exercise, the hardware setup needed should be as simple as possible. This avoids confusion among the participants, and could also possibly increase trust between inspected and inspecting state.

In the Darmstadt exercise, the students struggled with an additional provision: According to the agreement, the inspectors were allowed to test the scaler with a frequency generator in any way they thought useful. This freedom granted to the inspectors made it more difficult to cheat, because every possible action had to be anticipated.

Given the limitations of the exercise, which was first and foremost an educational experience, however, it remains to be seen if these insights can be transferred to credible inspection systems.

4 Tools for effective crowd sourcing

Carrying out actual exercises confirmed that the quality of the results not only depends on the participants (the "crowd") and their motivation, but also on the clarity of the problem proposed and the support provided by the organizers.

4.1 Exercise scenarios close to reality

While exercises can only cover very specific questions that are part of a broader process, it is important that the general scenario is as realistic as possible. Whenever changes from a possible real process are made, they should keep the scenario at least as consistent as possible. In particular, it is nearly impossible to carry out any exercise with real nuclear weapons, even radioactive sources and corresponding detectors pose a high burden on the organisers of a crowdsourcing exercise. Also other aspects might not be replacable in full detail (e.g. military safety regulations and personell).

For the course in Darmstadt, a simple device to simulate detector signals has been developed ("MeasureSim"). With "MeasureSim" it was possible to use realistic measurement electronics and a data acquisition system without the need for radioactive sources and detectors. The device is based on a RaspberryPi microcomputer and can simulate pulses (e.g. from spontanous fission events) on six different channels. The pulse rates can be easily selected using a web frontend. Pulses are generated with with randomized time periods between pulses to reflect the nature of radioactive decays. Although it has not been used in the above describe exercise, "MeasureSim" can also simulate signals that would be generated by taking a gamma spectrum of an item, e.g. a weapon-grade plutonium sample This is done using a Digital-Analog-Converter which generates output pulses with different voltage levels. The spectra have to be stored on the device as text files prior to the use.

In Princeton, measurements were carried out on a full-size mock-up reentry vehicle. Radioactive gamma sources were placed inside the mock-up and students used real detectors to acquire spectra. Students were given an "inspection toolkit" (cf. figure 1). Other than parts to build information barriers, it contained tags and seals with unique identification. In addition, intensive negotiations on the inspection protocol took place among the groups. All these points added to the students experience of a realistic simulation and "being an inspector".

4.2 Simple tools and open source advantages

Creativity of the crowd should not be impaired by complex prototype production, as this will slow down the process. The path from a new idea to its implementation needs to be as fast and seemlessly as possible.

For both student exercises, this was achieved by building systems around simple microprocessor platforms (e.g. RaspberryPi, BeagleBone, Arduino) that are cheap and provide easy ways to interact with the enviroment (e.g. sensors, LEDs, buttons). Because of the large existing user base, help files, documentation and many examples are available. In the two exercises described above, some participants had experiences with such tools. Any other tools for rapid prototyping are also helpful: printed circuit board production, 3D printers, tabletop CNC machines, laser cutter, multimeter and oscilloscope to troubleshoot ideas. These are common tools found in do-it-yourself development spaces such as Hackerspaces, FabLabs or TechShops (c.f. `http://www.fabfoundation.org/fab-labs/`).

Every microcontroller system and information barrier uses software. Quick results can be achieved if common programming languages (e.g. Python) can be used for these tools. It is important that comprehensive software libraries for all used tools (e.g. detectors, analog-digital-converter) are available for the prefered programming language.

Independent of the tool, quick and easy access can be achieved by focusing on the use of Open Source tools for hardware and software. It also has more advantages: After successful developments have been made, the free access increases reusibility of results for former participants but also for any other actor. Later, when ideas would be applied to actual problems, trust and transparency among states as well as among individual actors can increase, as every party can understand and comprehend the mechanisms behind a measurement system. Beyond transparency and trust, the open source approach can also help to increase participation [4].

Beyond physical tools, it is important to structure the teamwork and carefully document the innovation process. Interesting tools for efficient collaboration are online discussion software (e.g. web forum, wiki, mailinglist) or learning management systems, especially in university environments (e.g. Moodle). Groups should be advised to use revision control systems (e.g. git, svn). This helps both keeping track of changes and improving decision making process in the group.

5 Conclusion

In this paper, we presented the results of two crowdsourcing experiments applied to nuclear disarmament verification. Both exercises were carried in a university classroom environment and focused on the development of information barriers for nuclear weapons inspections.

The results show that engaging students not familiar with nuclear arms control can produce innovation and significant results. A good preparation of the exercise is necessary to faciliate and enhance the creative process. This includes realistic (or at least convincing) scenarios and the provision of rapid prototyping environments.

Future crowdsourcing exercises could differ in scope, size or targeted participants. Scaling up will require larger organizational structures and funding, but based on our experiences, it has the potential to lead to significant innovation.

References

[1] Keir Allen et al. "UK-Norway Initiative (UKNI) approach for the development of a Gamma Ray Attribute Measurement System with an integrated Information Barrier". In: *35th ESARDA Symposium proceedings, Bruges*. Ed. by F. Sevini. 2013.

[2] Daren C. Brabham. *Crowdsourcing*. Cambridge, Massachusetts: MIT Press, 2013.

[3] JL Fuller. *The functional Requirements and Design Basis for Information Barriers*. PNNL-13285. Pacific Northwestern National Laboratory, 1999.

[4] Moritz Kütt, Alexander Glaser, and Matthias Englert. "Open Source meets Nuclear Arms Control". In: *55th Annual INMM Meeting, Atlanta, GA*. 2014.

[5] Sébastien Philippe, Boaz Barak, and Alexander Glaser. "Designing Protocols for Nuclear Weapons Verification". In: *56th Annual INMM Meeting, Indian Wells, CA*. 2015.

[6] James Surowiecki. *The Wisdom of Crowds*. London: Abacus, 2005.

[7] Peter B. Zuhoski, Joseph P. Indusi, and Peter E. Vanier. *Building a Dedicated Information Barrier System for Warhead and Sensitive Item Verification*. BNL-66214. Brookhaven National Laboratory, 1999.