A Proposed Unified Framework for Techniques to Protect Sensitive Information

K.D. Jarman, W.K. Pitts, and A.J. Gilbert
Pacific Northwest National Laboratory, Richland, WA

July 30, 2015

Future arms control treaties must increasingly focus on verifying the presence or absence of an actual nuclear warhead or, in the case of dismantlement and disposition, components of warheads. The techniques to manage information in the context of inspection by a potential adversary are an active area of research. Funded by DOE/NNSA/Office of Defense Nonproliferation Research and Development (NA-22), PNNL has been evaluating the mathematical rigor of information protection by considering inspection actions as mathematical transforms. This approach differs from the traditional approach to Information Barriers (IBs) which has been in the language and context of vulnerability assessments.

As part of its work plan in this project, PNNL will be hosting a workshop to address this and other topics in information protection, inviting researchers to discuss how their approaches would solve a specific proposed verification problem with a common test item. A common analysis language and framework to evaluate proposed measurement techniques and information protection will be integral to this workshop, and can be further developed there. Developing the plan for this workshop is best done with the participation of a broad range of interested participants such as those expected to attend the INMM/ESARDA workshop to be held in Jackson Hole, WY on October 2-7, 2015. The expectation is that comments and feedback from those participants would be a key ingredient of the workshop plan that PNNL will submit to NNSA for its approval.

An example of rigorous evaluation is an approach developed and analyzed by PNNL using perceptual hash functions to obscure details of imaging information. This technique overlaps with image compression techniques developed in other areas. While simple to implement, the next stage of the evaluation revealed that for Arms Control, formal requirements led to a proof that a "dictionary attack" can be highly effective. Furthermore, the expert knowledge possessed by a treaty partner (expectations of component size, shape, and geometric arrangement) simplifies those attacks. The end result is that analyzing the image for a template match might require doing so in the context of full encryption such as a fully homomorphic encryption technique.

Based upon this work and evaluation of other techniques in areas such as medical imaging and compressive sensing, PNNL, together with project collaborators at Sandia National Laboratory and the University of Arizona, is evaluating the degree to which many proposed information protection techniques naturally fall into combinations of dimension reduction or sparse sampling. To date, these two broad categories seem to encompass many proposed techniques. The benefit of this approach is that it decouples the ultimate information protection of an approach, proven with mathematical rigor, from the practical details of implementation. If the most complex case of imaging information can be successfully characterized in this way, this approach might be useful to analyze the degree of protection in other, simpler data types.

Considering the breadth of approaches beyond those studied in this project, it is clear that all attempt to address one or more specific challenges that may arise in the context of future verification needs, including:

- increasing the number of attributes to counter cheating
- confirming spatial and related characteristics; e.g., using imaging
- allowing more information outside of IBs or eliminating components of IBs altogether
- reducing the complexity of IBs.

Furthermore, each approach may tend to focus solutions more on one end of the scale of securing information vs. confirming that the host or prover is not cheating. Not all approaches consider the full range of issues that arise, including operational implementation, and, in the case of templates, the "initialization problem" for the trusted item.

A common mathematical description and framework will help identify the problems solved and the challenges that remain, as well as the strengths and weaknesses of each approach. Our analysis of the perceptual hash was foundational in guiding our current research and understanding of similar approaches, strengthening the argument for mathematical formalism in this research area. In the specific area of research on zero-knowledge protocols, this need for common formalism and approach to proving technologies has also been raised in a recent session at the INMM Annual Meeting.

Common language will not only help researchers understand how to work together to eliminate the holes and strengthen technology. It will also help policy makers understand what options are available and how far each option can be pushed. Different technologies may be more or less appropriate depending on the actual negotiated verification specifics, which are in turn supported by knowledge of what the technologies can and cannot do. Understanding the options and their differences enables better informed decisions about which sets of technologies to further develop and test.

We propose that a framework should aim for sufficient generality to describe the core components of most, if not all, current and past measurement-based verification technology concepts (including both attribute and template approaches as special categories), in terms of the balance of sensitive information management and confident verification, and should focus on general characteristics of measurement systems, data transformations, and the mathematics of IBs to avoid getting mired in non-essential details of hardware and software.