Zero Knowledge Proofs and the Nuclear Deal

How to Ensure Iranian Compliance with the JCPOA

Katie Grosch

December 8, 2015

Contents

Abstract

Zero knowledge proofs are a tool used to confirm a statement, fact, or identity without revealing any information about it beyond its validity. This paper explores the applications of Zero Knowledge Proofs to physical properties, specifically nuclear disarmament. It presents a solution to the problem of verifying Iranian compliance with the Joint Comprehensive Plan of Action without violating Iranian sovereignty.

1 Introduction

Zero-knowledge proofs are a way to demonstrate the validity of a statement without revealing any information about the statement itself. It is well established that zero-knowledge proofs can be used to do things like verify solutions to an NP-hard decision problem, to enforce honest behavior in a game, and to prove identity [7]. Zero Knowledge Proofs can be used to ensure compliance with rules and requirements without requiring the prover to submit information about their practices. However, the applications of Zero Knowledge Proofs to the properties

of physical objects are not well studied.

One of the most hotly contested issues of the Iran nuclear deal, formally known as the Joint Comprehensive Plan of Action, is one of sovereignty. The demand of some nations that Iran open its borders to "anytime, anywhere" inspections [1] was decried by the mullahs as an unacceptable attack on Iranian sovereignty. Ultimately, the agreement settled on a complex plan for inspections. But could we have done better? This paper will evaluate the feasibility of applying Zero Knowledge Proof systems to enforce Iranian compliance with the agreement.

I will begin by explaining Zero Knowledge Proofs conceptually, as well as their current applications in the security field. I will look at what sort of problems can support a Zero Knowledge Proof, as well as how they prevent cheating by the prover. I will then delve into the idea of using Zero Knowledge Proofs to prove physical properties about objects, and look at the current state of research into nuclear verification through Zero Knowledge Proofs.

The focus will then shift to the Iran nuclear deal. I will explain the present terms of the agreement, as well as current enforcement

mechanisms. Then, I will present an alternative to the current verification system, using Physical Zero Knowledge proofs. This alternative would ensure compliance while also preserving Iranian sovereignty. I will look at the advantages and possible drawbacks of this mechanism, and conclude that the current system could be improved with a Zero Knowledge system.

2 To the Community

Zero Knowledge Proofs have, until very recently, been considered purely a digital concept, often more theoretical than useful. They are also relatively recent; they were first introduced in 1985 by Goldwasser, Micali, and Rackoff [8]. Similarly, information security is often confined purely to the world of computers, and its applications to geopolitics are ignored. I hope to show how we can integrate information security and international relations in a way that will leverage the best parts of each, and in a way that can be repeated over and over in different situations.

3 How Zero Knowledge Proofs Work

A zero knowledge proof is, at its most basic level, a proof that verifies something without revealing any information about that thing except that the property being proved is true. This premise seems contradictory on its face: how can a prover demonstrate a property of an object without revealing anything else about that object?

The proof is an iterative and interactive exchange [10]. In each iteration, the verifier presents the prover with an instance of the problem. The prover performs some private calculation, and returns the result to the verifier. The verifier then chooses to accept or reject.

The easiest way to get a feel for Zero Knowledge Proofs is through examples. Suppose that Alice wants to prove to Bob that she can tell how many gumballs are in a gumball machine, but she doesn't want to reveal the actual number. She begins by writing down her guess on a piece of paper. Then, while she can't see, Bob removes a few gumballs (as many as he wants). She then looks at the gumball machine again, and writes down how many gumballs are in the machine

now. Subtracting the second number from the first number, she can tell Bob how many gumballs he removed, and thus prove that she can tell how many gumballs are in the machine. Bob never actually knows how many gumballs there are, nor does he ever know how she calculated the total.

A key part of this type of proof is that the verifier is never one hundred percent certain of the proof. In the last example, Bob cannot be positive that Alice isn't merely an excellent guesser. But the more times the proof is repeated, the more confident he becomes. We can look at another example to see the math behind this idea. Suppose Alice is trying to prove to Bob that she has the password to a secret door hidden out of Bob's sight. In Figure 1 below, that puts Bob at position A.

Alice goes to the passage and selects which side of the door to go to, C or D (without Bob seeing). Bob then walks to point B and calls out which side of the door Alice should appear from. If she needs to go through the door, she quietly un-

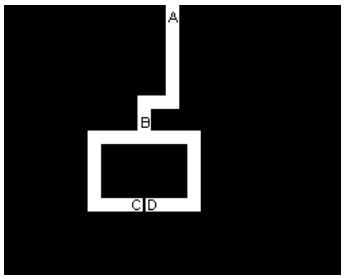

Figure 1

locks it and emerges from the other direction. If
she doesn't, she merely walks back out to Bob. At this point, Bob is only 50% confident that Alice actually knows the code to the door, because it's possible she just got lucky. But as they repeat the experiment again and again, his confidence increases: The chance that Alice doesn't know the code after i repetitions is

$$Probability\,(lucky\,guess) = 1/2^i \tag{1}$$

A Zero Knowledge Proof must be both sound and complete. It is sound if an honest Peggy will always convince an honest Bob. It is sound if cheating Peggy can convince honest Victor of a false statement with only a small probability.

4 Applications of ZKPs in Security

Zero knowledge proofs have many applications in authentication. Say Bob wants to verify that Peggy is who she says she is by proving that Peggy knows a secret that only Peggy would know. However, for obvious reasons, Peggy doesn't want to reveal her secret to Bob. This

situation gave rise to the Fiat-Shamir Identification Protocol [2], which serves as the bedrock of modern zero-knowledge identification protocols.

In fact, Zero Knowledge Proofs exist for all NP-hard problems, as long as a one-way function exists for the problem [7]. This can be proven through three-colorability or some other reduction [9]. The key is that in a Zero Knowledge Proof, the verifier learns no more about the solution, identity, or object than he knew to start, with the exception of the fact that the prover has a solution to the problem.

5 Physical Zero Knowledge Proofs

Physical Zero Knowledge Proofs are not a new idea. Several authors have described proofs of "physical" properties like the number of Pez in a dispenser or solving a Sudoku puzzle [9]. However, there has been little research into proof of properties that are solely physical, like chemical makeups, which cannot be solved using digital representations of the problem. A Physical Zero Knowledge Proof is a proof that an object

has a particular property, without revealing any information about the object besides whether it has that property. Physical Zero Knowledge Proofs will form the foundation of our verification protocol.

6 The Iran Nuclear Deal

Formally ratified earlier this year, the Joint Comprehensive Plan of Action, or JCPOA for short, is a deal negotiated between Iran and the P-5+1 nations that aims to prevent Iran from developing nuclear weapons, while simultaneously removing crippling sanctions on Iran's economy. The idea is that in return for strict limits on their nuclear capabilities, Iran will be allowed the opportunity to join the global economy.

However, Iran must also accept a wide array of inspections to their nuclear facilities. The International Atomic Energy Agency, or IAEA for short, will be in charge of those inspections. According to the agreement, the IAEA's inspections will include the use of satellite imagery, on-site searches for specialized equipment, document analysis,

interviews, and environmental sampling to check, for example, whether highly enriched uranium has been used at a site. The inspection team can also use ground-penetrating radar to search for hidden equipment.

Inspectors need to present evidence of nuclear activity before they can inspect a given site, which is short of the "anywhere, anytime" inspections demanded by some. Iran also has 24 days to comply with a request to inspect. Furthermore, Ayatollah Ali Khamenei, the supreme leader of Iran, has maintained that Iran's military facilities be free from surveillance, stating as recently as May, "We will not let foreigners inspect any military center" [4]. These and other restrictions were put into place to preserve Iranian sovereignty in the face of foreign inspections. But what if there were a way to inspect all facilities, anywhere, anytime, without foreign powers learning anything about Iran except that they are compliant with the JCPOA? This is where Zero Knowledge Proofs can help.

7 Current Verification System

A key problem of verifying nuclear disarmament is one of sovereignty. Often, nuclear facilities are highly classified, and governments are loath to allow international inspections any access, let alone unlimited access. Thus, the problem is as follows: how can we verify that a state is compliant with a nuclear treaty without any knowledge of their nuclear program besides its compliance? To date, this problem has been dealt with through complicated testing software that takes measurements and produces no output but a simple "yes/no" confirming or disconfirming compliance. The problems with this approach are immediately visible: how can both countries be sure that there is no hidden "trapdoor" in the software? How can they verify that the measurements aren't somehow being recorded, rather than merely verified? This system requires a lot of mutual trust and is difficult to implement [5].

8 A Better Alternative: ZKPs

For our proof of compliance, Iran will take on the role of the prover, and IAEA inspectors the role of the verifier. IAEA inspectors will take samples from nuclear facilities, with the goal of establishing that these samples do not contain unacceptable levels of bomb-grade material. There are two ways to create the material necessary to produce a bomb: either using centrifuges to pull out the U-235 from mined uranium, or using a nuclear reactor to transform mined uranium into plutonium [3]. Iran wants to ensure that these inspections do not reveal any information about Iran's nuclear program besides the fact that nuclear material is not weapons-grade.

This protocol will borrow heavily from Glaser, Barak, and Goldston's protocol for nuclear warhead verification [6]. In their protocol, the makeup of one or more nuclear warheads can be verified to be identical to a template without ever actually measuring the warhead's makeup. Instead, the prover pre-loads multiple detectors with the negative of the nuclear radiograph of the warheads. The inspector cannot access these preloaded radiographs. The inspector then pairs

up warheads and detectors, and ought to see that all the measurements are statistically similar. Dissimilar measurements would indicate that the warheads differ from each other. The inspector has access to what the chemical makeup of the warhead ought to be, and so can compare the images generated by the detectors and verify that they are all the same chemical makeup. If n warheads are being compared, then there is a $1/n$ chance that the inspector paired the correct negatives with the correct warheads. Repeated iterations of the protocol thus increase the inspector's confidence.

Our approach will use preloads of chemical makeups of nuclear samples from enrichment facilities. Inspectors can then verify that materials are not being enriched above 3.7% (the standard agreed to in the JCPOA) without ever actually measuring the radioactivity or chemical makeup of the samples. Essentially, Iran will supply the IAEA with samples from its nuclear facilities, claiming that none of the samples are enriched above 3.7%. Iran will also supply the IAEA with the negative of the radiograph loaded into a measurement device. The IAEA will not be able to access that negative. It will also have a tem-

plate sample and negative image of a non-weapons grade radiograph. The IAEA then assigns each sample to a detector and compares the combined images. With repeated iterations, the IAEA can increase its confidence that all of the samples are identical to a template sample and its negative.

Glaser et al. provide a more concrete example of how this works that makes it easier to understand. Here is that example, adapted to this situation. Imagine that the IAEA and Iran each have a bag of marbles. Iran wants to prove that it has x marbles, the same number of marbles as the IAEA, without letting the IAEA look at its marbles to see anything about their color, shape, or size. Iran prepares two bigger bags with *100-x* marbles in each. It then lets the IAEA dump each of their little bags of marbles into the bigger bags, one in each big bag. The IAEA then counts the number of marbles in both big bags and ensures that there are 100 marbles in each. In this example, if Iran cheats in preparing the buckets, then there is a *1/2* chance that the IAEA will catch the cheating.

In this case, the smaller bags of marbles represent Iran's nu-

clear samples and the IAEA's control samples. The larger bags are the radiograph negatives, and the number 100 is the template Iran's samples are being compared to. With this protocol, Iran's nuclear samples are tested for compliance, but crucially, the samples themselves are never tested by the IAEA. The IAEA just matches samples to radiograph negatives and counts the total result, as it were. The samples themselves are never measured, nor does any data about them leave Iran.

9 Potential Challenges

A frequent argument from the "anytime, anywhere" proponents of nuclear inspections is that no inspection can prevent the possibility that Iran will simply build a new nuclear facility in a location unknown to IAEA inspectors. If no inspection can be done at all, then it doesn't matter what methods are implemented in inspections; Iran could develop a nuclear weapon with impunity.

That isn't a problem that can be solved with a Zero Knowledge

Proof. This protocol isn't intended to prevent all "sneak out" capability by Iran, it's to implement an inspection strategy less susceptible to tampering and cheating at inspection time, and to create a framework in which countries with low levels of mutual trust can still verify each other's compliance with nuclear agreements.

Another potential issue with using Zero Knowledge Proofs to verify Iranian nuclear compliance is one of public relations. Explaining to the public that nuclear inspectors can verify Iranian compliance without knowing anything about their nuclear program is a tough sell at best. Furthermore, it's a fair assumption that states are reluctant to implement an inspection strategy that has the potential of providing less information about the state being inspected than current strategies. However, this concern can be dispelled with careful education about the increased reliability of a Zero Knowledge Protocol to verify compliance.

A problem with this is that inspectors can still learn about Iran's nuclear program simply by visiting nuclear sites (including military sites), which compromises the "zero knowledge" aspect of the

protocol. This problem is somewhat unavoidable, as inspectors need to have confidence in the authenticity of the samples they harvest. However, the level of knowledge they gain from mere sightseeing in the facilities is low.

A fourth problem is that ensuring compliance requires measurements on lots of different spectra: the number of reactors, the level of enrichment of nuclear material, and the quantity of their stockpile of nuclear material, to name a few. This protocol addresses a Zero Knowledge proof for only one of those: the level of enrichment of a given nuclear sample. However, the idea is not restricted to chemical sampling. This paper does not address the other measurements the IAEA will make, but it is certainly possible to conceive of Zero Knowledge Proofs that would measure the quantity of Iran's nuclear stockpile or the number of reactors it has without revealing either actual measurement to an outside party.

10 Conclusion

Information security and geopolitics are often considered at best only loosely connected, and there is little scholarship linking the two at present. This is a shame. The principles of verifying identities, proving properties of objects, and keeping hidden properties secret is almost impossible to unravel from international relations. Much like two users on a network, states in the international system are operating in a world where mutual trust is tenuous and information is precious. These two worlds have much in common, and the theoretical principles we use in one can also apply in the other.

This paper has shown that in theory, nuclear agreements can be enforced without taking measurements of materials that are classified by a given country. Applying this result to the Joint Comprehensive Plan of Action, it can be seen that ensuring Iranian compliance with the deal does not need to expose the chemical details of their nuclear program to the world. Their compliance can be measured without any actual measurements.

The ability to conduct these Zero Knowledge Proofs of physical properties will allow countries to better enforce compliance with nuclear treaties without losing control over classified information. It will increase international cooperation and trust, and allow for better control of nuclear weapons in the world. Zero Knowledge Proofs are the key to nuclear enforcement in the future.

References

[1] Critics Say U.S. Officials Promised 'Anytime, Anywhere' Inspections In Iran Nuclear Deal. In *All Things Considered*. August 2015.

[2] M. Abdalla, Jee Hea An, M. Bellare, and C. Namprempre. From Identification to Signatures Via the Fiat #x2013;Shamir Transform: Necessary and Sufficient Conditions for Security and Forward-Security. *IEEE Transactions on Information Theory*, 54(8):3631–3646, August 2008.

[3] William J. Broad and Sergio Peçanha. The Iran Nuclear Deal – A Simple Guide. *The New York Times*, March 2015.

[4] Thomas Erdbrink and David E. Sanger. Iran's Supreme Leader Rules Out Broad Nuclear Inspections. *The New York Times*, May 2015.

[5] James Fuller. Verification on the Road to Zero: Issues for Nuclear Warhead Dismantlement. *Arms Control Today*, 40(10):19–27, December 2010.

[6] Alexander Glaser, Boaz Barak, and Robert J. Goldston. A zero-knowledge protocol for nuclear warhead verification. *Nature*, 510(7506):497–502, June 2014.

[7] Oded Goldreich and Yair Oren. Definitions and properties of zero-knowledge proof systems. *Journal of Cryptology*, 7(1):1–32, December 1994.

[8] Shafi Goldwasser, Silvio Micali, and Charles Rackoff. The Knowledge Complexity of Interactive Proof Systems. *SIAM Journal on Computing*, 18(1):23, February 1989.

[9] Ronen Gradwohl, Moni Naor, Benny Pinkas, and Guy N. Rothblum. Cryptographic and Physical Zero-Knowledge Proof Systems for Solutions of Sudoku Puzzles. *Theory of Computing Systems*, 44(2):245–268, May 2008.

[10] Marinka Zitnik. Zero-knowledge Proofs. *XRDS*, 20(1):65–67, September 2013.