Ross Miller
Caroline Wright
Math 445 - Cryptography
April 16, 2012

An Introduction to Zero-Knowledge Proofs

Introduction

Zero-knowledge proofs are a type of interactive mathematical protocol that allow one party to demonstrate that they know some hidden information to another party, without revealing anything other than the fact that they have that knowledge. These two parties, she is often referred to as Peggy (the *prover*) and Victor (the *verifier*), respectively. Zero-knowledge proofs are useful as an authentication scheme between two parties when the identity of Victor, or the security of the communication channel, cannot be trusted. It can also be used for multiple parties to compute the solution to a function with private inputs, meaning no party knows all of the inputs, but each input is known by at least one of the parties. The classic example of this is the *millionaire problem*, wherein two millionaires wish to figure out who is richer without either wanting to reveal to anyone their personal wealth.

In most practical applications, zero-knowledge proofs are only probabilistically proofs. Therefore, when 100% confidence is not achievable through a certain implementation, it is more accurately referred to as zero-knowledge argument. When Peggy is trying to trick Victor, she is often referred to as Trudy. When this protocol is repeated, the chance that Trudy could manage to consistently fool Victor can be made arbitrarily small. Peggy's claim that she has this secret knowledge is known as the *statement*.

Every zero-knowledge protocol has three necessary qualities:

- *Completeness*: if Peggy adheres to the protocol, and the statement is true, then there is a 100% chance that Peggy will convince Victor.
- *Soundness*: if the statement is false, and the chance that Trudy can convince Victor of the statement's veracity can be made arbitrarily small.
- *Zero-Knowledge*: At the end of the protocol, Victor learns nothing about the statement other than it is in fact true with a certain degree of certainty. Victor can forge an honest looking transcript, and Peggy and Victor can conspire to make it look like Peggy knows the secret. Therefore, if an eavesdropper (Eve) has a complete transcript of all information exchanged during the protocol, she has learned nothing. Not even the veracity of the statement itself.

The Strange Cave of Ali Baba

The following is a visual story that demonstrates a proof-of-concept of zero-knowledge proofs. Ali Baba and his thieves have discovered a cave with a magic door. The door will only open if a secret word is spoken, but fortunately Peggy claims to know the secret word. Peggy does not

want to tell the secret unless she gets paid, and Ali Baba, who is acting as Victor, will not pay Peggy until he is convinced that she actually knows the secret. Ali Baba's cave has a fork in the road. The left and right directions both lead to opposite sides of the magic door.

(image source: Quisquater)

In this example, Peggy is already inside the cave at the fork in the road (on the right side of the image). Victor is standing outside the cave (at the top of the image) so that he cannot see Peggy. Peggy is then free to travel towards the magic door by either going left or right. Once she has made her choice, she sends a signal to Victor, who then enters the cave and waits at the fork in the road. Victor cannot see which direction Peggy has chosen, and Peggy is no longer able to change which direction she chose after Victor enters the cave, or else she will be seen, and Victor will know that she has violated the protocol. This is portion of the protocol is known as *bit commitment*.

Victor then randomly chooses "Left" or "Right" and calls out his selection to Peggy. In order to pass the protocol, Peggy must return to Victor from the same direction in the road that he chose. If Peggy went down the left side, and then Victor calls out "Right", she can only come out from the right side if she is able to speak the secret phrase and cross the door. However, if Peggy went down the same direction that Victor chose, she can turn around and head back to Victor without needing to pass through the magic door. Since a single iteration of this protocol is not very convincing, Victor can perform this protocol several times in a series, until he is satisfied with his level of confidence in whether or not Peggy knows the secret. Peggy must pass each iteration in order to pass the entire protocol. If she fails a single time to exit from the correct path, then Victor determines that she does not know the secret.

This protocol has completeness, because if Peggy knows the secret word she is able to pass freely through the door and come out through either path in the cave. It has soundness, because if Peggy does not know the secret, her only hope of fooling Victor is to anticipate each of his choices and go down the same direction that he is going to choose for that iteration. On average, she is going to make a mistake 50% of the time, unless she is somehow able to predict which side Ali Baba will choose every single time. Finally, this protocol is zero-knowledge only if Peggy is careful enough to utter the magic word softly enough so that Victor cannot hear it.

Fiat-Shamir Identification Scheme

Fiat-Shamir is a practical implementation of a zero-knowledge proof, with one caveat. Fiat-Shamir leaks one bit of information; specifically the sign bit of the secret, *s*. An adapted protocol called Feige-Fiat-Shamir fixes this information leak. Since the underlying concept is the same between the two, I will still only cover the inferior Fiat-Shamir, since it is more readily understandable. Unlike in The Strange Cave of Ali Baba, this implementation allows Victor to require Peggy to successfully perform many iterations of the protocol in parallel as well as in a series.

During this protocol, Peggy will demonstrate to Victor the validity of the statement, "Peggy knows a secret number, *s*, where $s^2 \equiv y \pmod{n}$, where *y* & *n* are public, *n* is the product of two large primes, and the factorization of *n* is not known by Peggy." Since *y* is the perfect square of $s \pmod{n}$, *y* is called a *quadratic residue modulo n*. This implementation relies on the assumed difficulty of finding quadratic residues modulo *n*, when *n* is the product of two large primes, and the factorization of *n* is unknown (Trappe 318).

Setup phase (this can be performed just once and used by many different provers):

- A trusted outside party produces a public number, *n*, by randomly selecting two large primes, *p* and *q*, and computing the product of those two numbers.
- Peggy selects a secret *s*, where *s* is co-prime to *n*.
- Peggy publicizes $s^2 \equiv y \pmod{n}$. In this case, *y*, is considered Peggy's public key

The protocol:

- Peggy chooses a random number, *r*, and sends $r^2 \pmod{n}$ to Victor
- Victor randomly sets $e = 0$ or $e = 1$, and sends *e* to Peggy
- Peggy computes $y \equiv r \cdot s^e \pmod{n}$, and sends *y* to Victor
- Victor checks if $y^2 \equiv r \cdot s^e \pmod{n}$. If it is, then Peggy has passed the iteration, otherwise, Victor rejects the entire proof. Victor repeats the protocol until he is satisfied.

At each iteration, Trudy only has a 50% chance of tricking Victor. If at any time, she provides him with an incorrect answer to one of his questions, then he determines that the statement is false, or in other words, she does not know the secret number.

This protocol satisfies the completeness property because if Peggy knows the secret, she will always be able calculate a valid *y*. If Peggy does not know the secret, then she can only fool Victor if she can anticipate his choice for *e*, which is limited by two choices. Therefore she has a 50% chance of fooling Victor by guessing alone. The *soundness error* is the percentage of confidence that Victor has of Peggy knowing the secret. In this case, the soundness error after one iteration is 50%. After *N* iterations, Trudy would only have a 2^{-N} chance of fooling Victor. As *N* approaches infinity, this ratio approaches 0%. Therefore, this protocol satisfies the soundness property of zero-knowledge proofs.

If Peggy gets lazy and uses the same random number, *r,* for multiple iterations, then the security of the system is compromised. If Victor can find two iterations where Peggy used the same random number *r*, where he chose $e = 0$ the first time and $e = 1$ the second time, then he knows both $y_0 \equiv r \pmod{n}$ and $y_1 \equiv r \cdot s \pmod{n}$. Victor can use the Extended Euclidian Algorithm to compute the modular multiplicative inverse y_0^{-1}. Then Victor can crack the secret by computing *s*

$\equiv y_0^{-1} \cdot y_1 \pmod{n}$. Once he thinks he may have found the secret, it can be easily verified by checking if $s^2 \equiv y \pmod{n}$.

If Trudy is sometimes able to successfully fool Victor, regardless of whether Victor chooses $e = 0$ or $e = 1$. She has to break the protocol to do so, and she will only be successful if she can correctly guess his random selection of e each and every time. Alternatively she would have to be able to actually solve the quadratic residue problem. This implementation relies on the assumption that quadratic residues are hard to compute. Therefore, Victor should allow Peggy a sufficiently long, but bounded time period within which she must respond with an answer.

Proving a Sudoku Solution with Zero-Knowledge

The paper, "Cryptographic and Physical Zero-Knowledge Proof Systems for Solutions of Sudoku Puzzles," presents various zero-knowledge protocols for "proving" that a particular Sudoku puzzle is solvable (Gradwohl). Once again, the two parties in this interactive proof are Peggy and Victor. Peggy and Victor both have the same starting Sudoku puzzle, where many of the squares are still empty.

The protocol:

1) Peggy turns the Sudoku puzzle over and copies the mirrored solution onto the back, so that each of the original numbers on the front face have the same number printed directly behind it.
2) Victor randomly tells Peggy one of three things: "Rows", "Columns", or "Subgrids". Peggy then grabs a pair of scissors and cuts the Sudoku puzzle into 9 rows, 9 columns, or 9 boxes, according to Victor's request. She then reconstructs the puzzle, face up, and shows to Victor that she has cut up the puzzle as requested. Since the puzzle is face up, she reveals zero information to Victor at this step.
3) Peggy then hides the segments from Victor's view again, and turns the segments face down so that the solution is showing. Peggy and Victor have prearranged a way to number the 9 segments. Peggy takes the first segment, and cuts that into 9 individual boxes. She scrambles these 9 boxes in her hand and then shows the back and front of each of these boxes to Victor. Since Victor has the uncompleted puzzle, he knows which numbers are already filled-in on the front side of the first segment. He verifies that the correct numbers are printed on the front, and that the numbers 1 through 9 have been filled in on the back, without duplicates or omissions. Whenever there is a number printed on the front, it must match the number written on the back.
4) Peggy repeats step 3 for segments 2 through 9.

It is clear that this protocol satisfies completeness. For soundness, observe that if Peggy could complete the protocol regardless of Victor's random request, then she must know the solution. Without knowing the solution, Peggy can construct a puzzle that will appear correct for two of the three possible requests Victor could make. Therefore the chance that a Trudy can fool Victor for each iteration is 1 in 3. Therefore the soundness error is 2/3. After k iterations, the soundness error is reduced to $(2/3)^k$.

A Zero Soundness Error Protocol with a Trusted Copier

If Peggy and Victor can agree on a trusted copier, it is possible to achieve completeness and 0 soundness error with a single iteration (Gradwohl). Peggy first performs step 1, then makes three double-sided copies. Victor asks for "Rows" on the first copy, "Columns" on the second, and "Subgrids" on the third. If Peggy passes each of these three protocols then there is a 0% chance that she can fool Victor.

Just as with the quadratic residue computation problem and the Sudoku solving problem, any problem that is in NP can be used to construct a zero-knowledge proof with minimal assumptions (Naor). Because of this fact, zero-knowledge proofs can be used in a limitless number of ways. Typically, zero-knowledge proofs are used in cryptography to make a party prove that they are behaving according to a protocol, without revealing any of the secret information of that particular protocol.

Works Cited

Gradwohl, Ronen; Naor, Moni; Pinkas, Benny; Rothblum, Guy N. (2007). "Cryptographic and Physical Zero-Knowledge Proof Systems for Solutions of Sudoku Puzzles".

Naor, Moni; Ostrovsky, Rafail; Venkatesan, Ramarathnam; Yung, Moti. (1998). "Perfect Zero-Knowledge Arguments for NP Using any One-Way Permutation". Journal of Cryptology 11:87-108.

Quisquater, Jean-Jacques; Guillou, Louis C.; Berson, Thomas A. (1990). "How to Explain Zero-Knowledge Protocols to Your Children". *Advances in Cryptology - CRYPTO '89: Proceedings* **435**: 628–631.

Trappe, Wade, and Lawrence C. Washington. *Introduction to Cryptography: with Coding Theory*. 2nd ed. Upper Saddle River, N.J.: Pearson Prentice Hall, 2006. Print.