CS 2051: Honors Discrete Mathematics
Spring 2023 Homework 7 Supplement outline

Sarthak Mohanty

Title: ECC is the new RSA

Overview

Elliptic Curve Cryptography (ECC) is one of the most powerful cryptosystems in use today. Companies are using ECC everything to secure everything from our customers' HTTPS connections to how we pass data between our data centers. In fact, based on currently understood mathematics, ECC provides a significantly more secure foundation than first generation public key cryptography systems like RSA, for reasons we'll explore in detail later. The specific implementations you might learn in this supplement aren't as important as the general framework, understanding, and ideas you'll develop.

Cryptography Primer

Symmetric vs Asymmetric Systems

In cryptography, there are two main types of encryption algorithms: symmetric and asymmetric. Symmetric encryption uses the same key to both encrypt and decrypt the message, while asymmetric encryption uses different keys for encryption and decryption.
Symmetric encryption algorithms are generally faster and more efficient, but they require the sender and the recipient to share a secret key. The most prominent example is known as AES[1]
Asymmetric, or public-key encryption algorithms, on the other hand, are slower and less efficient, but they allow the sender to send a message securely to the recipient without sharing a secret key. Examples include RSA, Diffie-Helman, and ECCDH.

Trapdoor Functions: RSA

Trapdoor functions are an essential component of asymmetric encryption algorithms. A trapdoor function is a function that is easy to compute in one direction, but difficult to compute in the opposite direction without knowledge of additional information, such as a secret key.

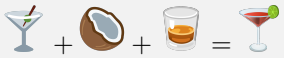
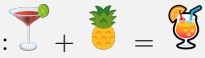
As an example, the trapdoor function in RSA encryption is the modular exponentiation function, $c = m^e \mod N$, where m is the plaintext, e is the public exponent, N is the modulus and c is the ciphertext. The trapdoor information is the private key, which consists of the factors of N. In the opposite direction, decryption is performed by computing $m = c^d \mod N$, where d is the private exponent derived from the private key.

[1] some of your classmates have chosen this as their group project!

Part 1: Diffie-Helman

Diffie-Hellman Key Exchange

The Diffie-Hellman key exchange is a cryptographic protocol that allows two parties to generate a shared secret key over an insecure communication channel. The protocol is named after its inventors, Whitfield Diffie and Martin Hellman.
There is a classic story of how diffie-helman (and in general) works, of which a modified version is below:

> Professor Brito and Lianet (Brito's wife!) have a famous Piña Colada recipe that they keep secret from the rest of the world. It is so secret in fact, that neither Brito nor Lianet know the full parts of the recipe! However, using their knowledge on how the Diffie-Helman key exchange algorithm works, they have devised a unique way of making the drink.
>
> To make their famous Piña Colada, Brito and Lianet have their own secret ingredients, which they mix separately. Brito mixes a secret amount of rum to a glass, while Lianet mixes a secret amount of pineapple juice to another glass. These individual mixtures represent the private keys of each party:
>
> Brito:  Lianet: 
>
> Next, they swap glasses. This represents the public key exchange. Note that because the ingredients were mixed in, it is very difficult for either Brito or Lianet to determine the exact contents of the current glass.
>
> Brito:  Lianet: 
>
> Once they have exchanged their public keys, they each use their own private key to mix with the other party's public key to generate a shared secret key. Finally, they come together and mix the secret amount of the remaining ingredients in a third glass.
>
> Brito:  Lianet: 
>
> This shared secret key can be used to establish a secure communication channel, just like the delicious Piña Colada they have created together.

In this analogy, the rum and pineapple juice represent the private keys, while the remaining ingredients represent the public key exchange. The shared secret key is the final product, which can be used for secure communication between the two parties.

See the textbook for actual diffie-helman.

Trapdoor Function: Discrete Log Problem

Attacks

Brute Force

All public key schemes are in theory susceptible to a brute force attack. Try all possible values. This takes $\mathcal{O}(n)$ time.

Baby-Step Giant-Step algorithm

This algorithm trades off space to improve the running time over brute force. The idea is to write $x = i\lceil n \rceil + j$ for $0 \leq i, j \leq \lceil n \rceil$. For simplicity let $m = \lceil n \rceil$. First compute g^j for all $0 \leq j \leq m$. Then we compute $y \cdot g^{-im}$ for each $0 \leq i \leq m$. Notice that both can be done in $O(\sqrt{n})$ time. If we could find a collision $y \cdot g^{-im} = g^j$, then $x = im + j$. To find this collision efficiently we store all g^j (for all $0 \leq j \leq m$) by hash-sets and when we compute $y \cdot g^{-im}$ we simply check if that value is already present in the set. This finds the collision in $O(\sqrt{n})$ time (as we need to check if the value is present in the set only $O(\sqrt{n})$ times). Overall this algorithm uses $O(\sqrt{n})$ space and time.

General Number Field Sieve (GNFS)

The General Number Field Sieve is a factorization algorithm that is particularly efficient for numbers with large prime factors. It is a multi-step process that involves finding a smooth polynomial and then finding the roots of that polynomial.

It first selects a sequence of integers that has a high probability of containing a factor of the number being factored. It then tries to find smooth numbers in this sequence, which are numbers whose prime factors are all below a certain bound. Finally, it uses linear algebra to combine these smooth numbers in a way that allows it to extract non-trivial factors of the number being factored. GNFS is one of the most efficient known factorization algorithms for large integers.

There is one final important note about factoring algorithms such as the general number field sieve. Such algorithms get more efficient as the size of the numbers being factored get larger. The gap between the difficulty of factoring large numbers and multiplying large numbers is shrinking as the number (i.e. the key's bit length) gets larger. As the resources available to decrypt numbers increase, the size of the keys need to grow even faster. This is not a sustainable situation for mobile and low-powered devices that have limited computational power. The gap between factoring and multiplying is not sustainable in the long term.

All this means is that RSA is not the ideal system for the future of cryptography. In an ideal Trapdoor Function, the easy way and the hard way get harder at the same rate with respect to the size of the numbers in question. We need a public key system based on a better Trapdoor. In 1985, such a system was found, revolving around an (then) arcane field known as elliptic curves.

In this part, you'll fully implement the Diffie-Helman key exchange and try to break it on small numbers.
Class Actor:

- `computePublicKey(self)`: See docs for details
- `computeSecret(self, offer)`: See docs for details

Class BadActor

- `brute(self, g, y, n)`:: This function takes in the parameters for the discrete log problem, and returns the correct answer using brute force.
- `bsgs(self, g, y, p)`: Same thing as above, but with baby-step giant-step algorithm.

Part 2: Enter Elliptic Curves

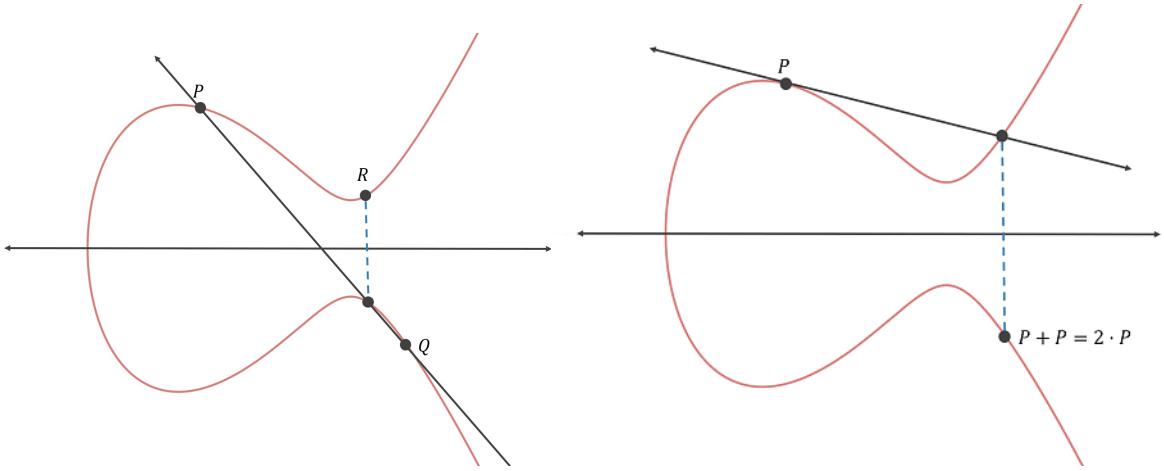
The standard form for an elliptic curve is

$$y^2 = x^3 + Ax + B,$$

where A and B are constants. This will be referred to as the Weierstrass equation for an elliptic curve. For reasons that will be discussed later, also include the point $\{\infty\}$ in this function.

The Group Law

One important aspect of elliptic curves is that we can start with two points, or even one point on an elliptic curve, and produce another point.



(a) Point Addition (b) Point scalar multiplication

Figure 1: Group law visualized.

In this part, you'll implement point addition and point multiplication on elliptic curves
In this part, you'll implement the following functions:

- `point_addition(elliptic_curve, P, Q)`: This function takes in an elliptic curve
- `point_scalar_multiplication(elliptic_curve, k, P))`: This function takes in an elliptic curve of the form described, a scalar integer k, and a point P, and repeatedly multiplies this result to output a number.

You can run `visualize_addition` and `visualize_multiplication` to generate the corresponding visualizations mentioned above. Notes: This part is just algebra and geometry. Since I know it's been a while since many of you touched this, here's a few steps you get you started.

- For point addition, we assumed that $P \neq Q \neq \infty$. For this reason, we can draw the line L through P_1 and P_2. Its slope is
$$m = \frac{y_2 - y_1}{x_2 - x_1}.$$
If $x_1 = x_2$, then L is vertical, so the resulting point is just ∞. Now let's assume that $x_1 \neq x_2$. The equation of L is then
$$y = m(x - x_1) + y_2.$$
To find the intersection of this line with the elliptic curve, we can substitute the above equation into the formula for an elliptic curve
$$y^2 = x^3 + ax + b.$$
Finally, we can reflect the above point across the x-axis to obtain the point $P_3 = (x_3, y_3)$.
- For point scalar multiplication, we have that $P = Q = (x_1, y_1)$. When two points on a curve are very close to each other, the line through them approximates a tangent line. Therefore, when the two points coincide, we take the line L through them to be the tangent tine. Implicit differentiation allows us to find the slope m of L:
$$2y\frac{dy}{dx} = 3x^2 + A, \quad \text{so} \quad m = \frac{dy}{dx}\frac{3x_1^2 + A}{2y_1}.$$
If $y_1 = 0$ then the line is vertical and we set $P_1 + P_2 = \infty$, as before. Otherwise, we plug in values and solve as before.

Part 3: In The Field

The elliptic curve cryptography (ECC) uses elliptic curves over the finite field $\mathbb{F}_p$ (where p is prime and $p > 3$). For example, the "Bitcoin curve" secp256k1 takes the form:

$$y^2 \equiv x^3 + 7 \pmod{p}$$

Use this for part 3

In this part, you'll visualize elliptic curves over finite fields, as well as implement point addition and point multiplication over these fields

In this part, you'll implement the following functions:

- `generate_point_cloud(a, b, p)`
- `point_addition(P, Q, a, b, p)`: This function takes in an elliptic curve and two points, and adds them
- `point_scalar_multiplication(P, n, a, b, p)`: This function takes in an elliptic curve of the form described, a scalar integer k, and a point P, and repeatedly multiplies this result to output a number.

Part 4: ECC in Action: Elliptic-curve Diffie Helman

The Elliptic-curve Discrete Log Function

Trapdoor function in Elliptic Curve Cryptography (ECC). In ECC, the trapdoor function is the scalar multiplication operation $Q = kP$, where P is a point on the elliptic curve, k is the scalar and Q is the result of the scalar multiplication. The trapdoor information is the private key, which is the scalar k. Decryption is performed by computing $Q = -kP$.

For carefully chosen (by cryptographers) finite fields and elliptic curves, the ECDLP problem has no efficient solution.

The multiplication of elliptic curve points in the group $\mathbb{F}_p$ is similar to exponentiation of integers in the group $\mathbb{Z}_p$ (this is known as multiplicative notation) and this is how the ECDLP problem is similar to the DLP problem (discrete logarithm problem).

See this video for implementation details

In this part, you'll implement a elliptic-curve diffie helman key exchange and try to break it on small numbers. You should directly call the methods you constructed in Part 2. You should replicate the methods you created in Part 1, this time accommodating for the elliptic-curve element.

Class Actor:

- `computePublicKey(self)`: See docs for details
- `computeSecret(self, offer)`: See docs for details

Class BadActor

- `brute(self, g, y, n)`:: This function takes in the parameters for the discrete log problem, and returns the correct answer using brute force.
- `bsgs(self, g, y, p)`: Same thing as above, but with baby-step giant-step algorithm.

Conclusion

RSA vs ECC

Since RSA and ECC essentially serve the same purpose, there have been many heated discussions as to which implentation to use. The main pro of using ECC is the fact that it uses fewer memory and CPU resources, important as mobile computing becomes more ubiquitous. However, there are a few flaws with ECC. First,

there are only a few curves that work and the NIST is in contorl of most of them, but people distrust NIST. Second,

Finally, neither ECC nor RSA are secure against quantum computers.

Real World Application

Even with all those downsides, more and more companies are using ECC, including

Amazon, Bing, Dropbox, Facebook, Flickr, GitHub, Instagram,
LinkedIn, MSN, Netflix, Pinterest, PirateBay, Quora, Snapchat,
SoundCloud, Spotify, StackOverflow, Tumblr, Twitter, Uber,
Vimeo, Vine, Yahoo, Yelp, YouTube, Wikipedia, Wordpress,...

Two prominent examples:

- ChatGPT
- ProtonVPN

to do: explain how you can see ecc implentation on website. Click on lock on top left of website, click on "connection is secure", click on "certificate is valid", you should see ECC used for certificate
assymetric not often used since takes long time. Instead, key exchange is done using assymetric andthen symmetric is used to send the actual messages, as can be seen here

Submission Instructions (10 pts)

After you fill the appropriate functions, submit the following files to Gradescope and make sure you pass all test cases:

- `diffie_helman.py`
- `elliptic_curves.py`
- `ECDH.py`

Notes

- The autograder may not reflect your final grade on the assignment. We reserve the right to run additional tests during grading.
- Do not import additional packages, as your submission may not pass the test cases or manual review.

References

explanation for group law name is here
why infinity is added here
Textbook for elliptic curves is here