Chapter 13

Security & Cryptography

Every day people order and pay online – but is it secure?

13.1 Transport Layer Security

Remarks:

- Let's assume that Alice and Bob want to communicate with each other over the internet, but they are afraid of a potential eavesdropper Eve who might be able to intercept any communication between Alice and Bob.
- Alice and Bob don't want Eve to be able to read their messages. Therefore, they encrypt their messages using a bulk encryption algorithm (see Section 13.2).
- For the encryption algorithm, they need to agree on a secret key using a key exchange protocol (see Section 13.3).
- When Alice receives a message, how can she be sure that the message hasn't been modified on the way from Bob to her? Alice and Bob use a message authentication algorithm (see Section 13.4) to ensure integrity of the communication.
- Let's assume that Alice hasn't met Bob in person before. How can she be sure that she is really communicating with Bob and not with Eve? She would ask Bob to authenticate himself (see Section 13.5).
- The Transport Layer Security (TLS) protocol is a standardized protocol which offers all of these features.

Protocol 13.1 (Transport Layer Security, TLS)**.** *TLS is a network protocol in which a client and a server exchange information in order to communicate in a secure way. Common features include a bulk encryption algorithm, a key exchange protocol, a message authentication algorithm, and lastly, the authentication of the server to the client.*

Remarks:

- TLS is the successor of Secure Sockets Layer (SSL). However, sometimes in practice the term SSL includes (the newer) TLS as well.
- HTTPS (Hypertext Transfer Protocol Secure) is not a protocol on its own, but rather denotes the usage of HTTP via TLS or SSL.
- SSH (Secure Shell), even though close in name to SSL, is something different: It is a protocol to allow a client to remotely access a server, e.g., for a command-line interface.

13.2 Secret Sharing & Bulk Encryption

"*Three may keep a secret, if two of them are dead.*" – Benjamin Franklin

Definition 13.2 (Perfect Secrecy)**.** *An encryption algorithm has perfect secrecy, if the encrypted message reveals no information to an attacker, except for the possible maximum length of the message.*

Definition 13.3 (Threshold Secret Sharing)**.** *Let $t, n \in \mathbb{N}$ with $1 \leq t \leq n$. An algorithm that distributes a secret among n participants such that t participants need to collaborate to recover the secret is called a (t,n)-threshold secret sharing scheme.*

Algorithm 13.4 (n, n)-Threshold Secret Sharing

Input: A secret k, encoded in binary representation of length $l(k)$.

Secret distribution

1: Generate $n-1$ random binary numbers k_i of length $l(k)$ and distribute them among $n-1$ participants
2: Give participant n the value k_n as the result of *XOR* of k and $k_1, \ldots, k_{n-1}$, i.e., $k_n = k \oplus k_1 \oplus k_2 \oplus \cdots \oplus k_{n-1}$

Secret recovery

1: Collect all n values $k_1, \ldots, k_n$ and obtain $k = k_1 \oplus k_2 \oplus \cdots \oplus k_{n-1} \oplus k_n$

Theorem 13.5. *Algorithm 13.4 has perfect secrecy even if $n-1$ participants collaborate.*

Proof. The theorem holds as applying the *XOR* operation $\oplus$ to a random bitstring and k results in a random bitstring. □

Remarks:

- How can we achieve a (t, n)-threshold secret sharing scheme with perfect secrecy?
- At this point, let us introduce the relevant notation for this chapter. We will use k for keys, m for messages, p for primes, g for primitive roots, and c for ciphertext (encrypted messages). Generally speaking,

an encryption algorithm encrypts a plain message m by applying a key k, resulting in ciphertext c.

Algorithm 13.6 (t, n)-Threshold Secret Sharing

Input: A secret k, represented as a real number.

Secret distribution

1: Generate $t-1$ random $a_1, \ldots, a_{t-1} \in \mathbb{R}$
2: Obtain a polynomial f of degree $t-1$ with $f(x) = k + a_1 x + \cdots + a_{t-1} x^{t-1}$
3: Generate n distinct $x_1, \ldots, x_n \in \mathbb{R} \setminus \{0\}$
4: Distribute $(x_1, f(x_1))$ to participant P_1, …, $(x_n, f(x_n))$ to P_n

Secret recovery

1: Collect t pairs $(x_i, f(x_i))$ from at least t participants
2: Use Lagrange's interpolation formula to obtain $f(0) = k$

Remarks:

- With at most $t-1$ pairs $(x_i, f(x_i))$, there are infinitely many possible polynomials with different values for $f(0)$.
- There are many other (t, n)-threshold secret sharing schemes, e.g., with intersecting hyperplanes.
- Note that in practice, a finite field of prime order instead of real numbers is used.
- How can two parties communicate securely and privately?

Algorithm 13.7 One-Time Pad

Input: A message m known to Alice, and a symmetric key k (as a random bitstring) of length $l(k)$, known by both Alice and Bob.

Encryption

1: Alice sends $c = m \oplus k$ to Bob

Decryption

1: Bob obtains m by $m = c \oplus k$

Corollary 13.8. *Algorithm 13.7 has perfect secrecy.*

Remarks:

- In crypto, it's always Alice and Bob, with a possible attacker Eve.
- Algorithm 13.7 only works if the message m has the same length as the key k.
- How can we encrypt messages of arbitrary length?

Definition 13.9 (Bulk Encryption Algorithm)**.** *A bulk encryption algorithm can securely encrypt a message of any size.*

Definition 13.10 (Electronic Code Book, ECB)**.** *Given a method to encrypt a block of x bits, ECB encrypts a message of length rx by splitting the message into r blocks of length x, encrypting each block separately.*

Remarks:

- Do we now have a secure method to easily encrypt a large message, if we can encrypt small blocks, each using the same one-time pad?
- Suppose you have two message blocks m_1, m_2 of the same length, encrypted with k, resulting in c_1, c_2. However, you can obtain $m_1 \oplus m_2 = c_1 \oplus c_2$, giving you information about m_1 and m_2.

Definition 13.11 (Cipher Block Chaining, CBC)**.** *Given a method f to encrypt a block of x bits, CBC encrypts a message of length rx by splitting the message into r blocks $m_1, m_2, \ldots, m_r$, each of length x, encrypting (the plaintext of) each block XORed with the previous encrypted block, i.e., $c_i = f(m_i \oplus c_{i-1})$. The first block c_0 is initialized randomly.*

Remarks:

- Are we secure now? Using the same technique as in the last remark, you can again get, e.g., $m_4 \oplus m_5$.
- CBC is still one of the standard techniques though when encrypting blocks successively, as more advanced algorithms are not susceptible to this simple attack for one-time pads. An example would be the advanced encryption standard (AES). Using AES with CBC is an example of a bulk encryption algorithm. The operation of AES is beyond the scope of this short chapter however.
- Note that Algorithm 13.7 has one big disadvantage – Alice and Bob need to agree on a large random number first! While this is feasible for, e.g., secret agents, it is quite impractical for everyday usage.

13.3 Key Exchange

How to agree on a common secret key in public, if you never met before?

Definition 13.12 (Primitive Root)**.** *Let $p \in \mathbb{N}$ be a prime. $g \in \mathbb{N}$ is a primitive root of p if the following holds: For every $h \in \mathbb{N}$, with $1 \leq h < p$, there is a $k \in \mathbb{N}$ s.t. $g^k = h \mod p$.*

Example 13.13. *$g = 2$ is a primitive root of $p = 5$, because $2^1 = 2 \mod 5$, $2^2 = 4 \mod 5$, $2^3 = 3 \mod 5$, and $2^4 = 1 \mod 5$. There exists one more primitive root of 5.*

Algorithm 13.14 Diffie-Hellman Key Exchange

Input: Publicly known prime p and a primitive root g of p.
Result: Alice and Bob agree on a common secret key.

1: Alice picks a secret key $k_A \in \{1, 2, \ldots, p-1\}$ and sends $A = g^{k_A} \mod p$ to Bob.
2: Bob picks a secret key $k_B \in \{1, 2, \ldots, p-1\}$ and sends $B = g^{k_B} \mod p$ to Alice.
3: Alice calculates $K = B^{k_A} \mod p$
4: Bob calculates $K = A^{k_B} \mod p$

Example 13.15 (Algorithm 13.14 with $p = 5$ and $g = 2$)**.** *Let's assume that Alice picks $k_A = 2$ and Bob picks $k_B = 3$. Thus, Bob receives $A = 2^2 \mod 5 = 4$ and Alice receives $B = 2^3 \mod 5 = 3$. Then, Bob calculates $4^3 \mod 5 = 4$, and Alice calculates $3^2 \mod 5 = 4$. Hence, Alice and Bob have agreed on the common secret key of 4.*

Theorem 13.16. *In Algorithm 13.14, Alice and Bob agree on the same key K.*

Proof. Everything mod p, we have

$$K = B^{k_A} = \left(g^{k_B}\right)^{k_A} = g^{k_B k_A} = g^{k_A k_B} = \left(g^{k_A}\right)^{k_B} = A^{k_B} = K \mod p.$$

□

Remarks:

- There are sophisticated methods to quickly find primitive roots, but they are beyond the material covered in this chapter.
- How secure is Algorithm 13.14?

Definition 13.17 (Discrete Logarithm Problem)**.** *Let $p \in \mathbb{N}$ be a prime, and let $g, a \in \mathbb{N}$ with $1 \leq g, a < p$. The discrete logarithm problem is defined as finding an $x \in \mathbb{N}$ with $g^x = a \mod p$.*

Remarks:

- Intuitively, the best approach to calculate the common secret key of Algorithm 13.14 from the publicly known p, g, g^{k_A}, g^{k_B} is to solve the discrete logarithm problem. This is also the best known attack.
- However, for some classes of primes there are better attacks, which is why one often resorts to so-called safe primes p, where $p' = (p-1)/2$ is also a prime.
- How to find big enough primes though? Deterministic methods are still too slow in practice. Thus, let's go probabilistic with the following primality test.

Algorithm 13.18 Probabilistic Primality Testing

Input: An odd number $p \in \mathbb{N}$.
Result: Is p a prime?

1: Let $j, r \in \mathbb{N}$ and j odd with $p - 1 = 2^r j$
2: Select $x \in \mathbb{N}$ uniformly at random, $1 \leq x < p$
3: Set $x_0 = x^j \mod p$
4: **if** $x_0 = 1$ or $x_0 = p - 1$ **then**
5: Output "p `is probably prime`" and **stop**
6: **end if**
7: **for** $i = 1, \ldots r - 1$ **do**
8: Set $x_i = x_{i-1}^2 \mod p$
9: **if** $x_i = p - 1$ **then**
10: Output "p `is probably prime`" and **stop**
11: **end if**
12: **end for**
13: Output "p `is not prime`"

Lemma 13.19. *Algorithm 13.18 is correct with probability* 75% *if it outputs "p is probably prime", and* 100% *correct if it outputs "p is not prime".*

Corollary 13.20. *The runtime of Algorithm 13.18 is* $O(r) \in O(\log p)$

Remarks:

- The proof for the probabilistic correctness of the primality test in Algorithm 13.18 goes beyond the material covered in this lecture.
- Algorithm 13.18 is a Monte Carlo algorithm as its (fast) runtime is deterministic, but the output can be wrong with bounded probability. However, running the algorithm again on the same p, but with different x, produces an independent result, allowing to bound the error probability by $\frac{1}{4^r}$ in r runs.
- A simple method to find big primes is thus as follows: Pick a big random number p, with p being odd. Run Algorithm 13.18

  until p is prime with the desired probability of $1 - \varepsilon$. If p is not prime, pick another p. According to the prime number theorem, the average distance between two primes of size at most n is just $\ln n$, i.e., there is a good chance to find a big prime.
- While it is easy to find big primes, the problem of factorization is believed to be hard: I.e., given some integer x, find the prime factors of x. Many cryptographic protocols rely on the (perceived) hardness of the factorization problem, most famously RSA.

Definition 13.21 (Man in the Middle Attack)**.** *A man in the middle attack is defined as an attacker Eve deciphering or changing the messages between Alice and Bob, while Alice and Bob believe they are communicating directly with each other.*

Theorem 13.22. *The Diffie-Hellman Key Exchange from Algorithm 13.14 is vulnerable to a man in the middle attack.*

Proof. Assume that Eve can intercept and relay all messages between Alice and Bob. That alone does not make it a man in the middle attack, Eve needs to be able to decipher or change messages without Alice or Bob noticing. However, Eve can emulate Alice's and Bob's behavior to each other, by picking her own k'_A, k'_B, and then agreeing on common keys $g^{k_A k'_B}$, $g^{k_B k'_A}$ with Alice and Bob, respectively. Thus, Eve can relay all messages between Alice and Bob while deciphering and (possibly) changing them, while Alice and Bob believe they are securely communicating with each other. □

Remarks:

- It is a bit like concurrently playing chess with two grandmasters: If you play white and black respectively, you can essentially let them play against each other by relaying their moves.
- How do we fix this? One idea is to personally meet in private first, exchange a common secret key $k_{A,B}$, and then use this key for secure communication. Now a man in the middle cannot change the key.

Definition 13.23 (Forward Secrecy)**.** *A sequence of secured communication rounds has the property of forward secrecy, if discovering the secret key(s) of a single communication round does not reveal the content of past communication rounds.*

Remarks:

- So Alice and Bob cannot use the same secret key multiple times.

Algorithm 13.24 Diffie-Hellman Key Exchange with Forward Secrecy

---

Input: Alice's and Bob's common secret key $k_{A,B}$, and furthermore a prime p with a primitive root g for p.
Result: A Diffie-Hellman key exchange not vulnerable to a man in the middle attack, and with forward secrecy.

1: Bob picks a random number $k_B \in \{1, 2, \ldots, p-1\}$ and sends Alice $\left(g^{k_B} \mod p\right)$ encrypted with $k_{A,B}$ as c_B as a challenge
2: Alice picks a random number $k_A \in \{1, 2, \ldots, p-1\}$ and sends $\left(g^{k_A} \mod p\right)$ encrypted with $k_{A,B}$ as c_A to Bob as a challenge
3: Alice and Bob decrypt the respective messages, and Alice sends $g^{k_B} + 1$ encrypted with $k_{A,B}$ to Bob as a response (and Bob as well with $g^{k_A} + 1$)
4: If decryption yields $g^{k_A} + 1$ for Alice, and $g^{k_B} + 1$ for Bob, respectively, they accept the round key $g^{k_A k_B} \mod p$

Lemma 13.25. *Algorithm 13.24 has the property of forward secrecy and is not vulnerable to a man in the middle attack, if encryption with $k_{A,B}$ is secure.*

Proof. For a man in the middle attack, Eve needs to be able to decrypt and encrypt with $k_{A,B}$ to convince Alice and Bob that they directly communicated with each other, which is a contradiction to the security assumption.

Regarding forward secrecy, if the attacker Eve gathers the secret key $g^{k_A k_B}$ of a communication round, she can decrypt the messages of this communication round.

Even if Eve gains access to $k_{A,B}$, she cannot gain access to the keys generated in past communication rounds. □

Remarks:

- Observe that forward secrecy only applies to communication rounds in the past. If Eve gains access to $k_{A,B}$, she can perform man in the middle attacks in future communication rounds.
- However, we have a new inconvenience: Alice and Bob need to agree on a secret key $k_{A,B}$ beforehand. Furthermore, with n participants, everyone needs $n-1$ different keys.

13.4 Message Authentication & Passwords

"*I've been imitated so well I've heard people copy my mistakes.*" – Jimi Hendrix

Definition 13.26 (Replay Attack)**.** *In a replay attack a previously valid message from Alice to Bob is sent again from an eavesdropper Eve to Bob.*

Remarks:

- An easy way to prevent replay attacks is to include time stamps in messages. Bob can detect a replay attack, if the time stamp is too old or multiple messages with the same time stamp arrive. Another idea is to use *nonces* (numbers only used once), with the sender and receiver keeping track of the nonces used so far.
- Another issue is that an attacker could change an encrypted message without knowing the content

Definition 13.27 (Malleability)**.** *If ciphertext c can be changed to c' such that the receiver decrypts it into a different message m' without noticing, the encryption algorithm is malleable.*

Remarks:

- Thus, we need a way to ensure that the messages cannot be changed by an attacker. A natural solution are hash functions. However the hash functions described in Chapter 8 are not secure.

Definition 13.28 (One-Way Hash Function)**.** *A hash function $h : U \to S$ is called one-way, if for a given $z \in S$ it is computationally hard to find an element $x \in U$ with $h(x) = z$.*

Definition 13.29 (Collision Resistant Hash Function)**.** *A hash function $h : U \to S$ is called collision resistant, if it is computationally hard to find elements $x \neq y$, $x, y \in U$, with $h(x) = h(y) \in S$.*

Remarks:

- It can be shown that a collision resistant hash function is also a one-way hash function.

Theorem 13.30 (Example for a Collision Resistant Hash Function)**.** *Let* $p = 2q + 1$ *be a safe prime, with primitive roots* $g_1 \neq g_2$ *of* p*. The hash function* $h : \{0, \ldots, q-1\} \times \{0, \ldots, q-1\} \to \mathbb{Z} \setminus \{0\}$ *with* $h(x_1, x_2) = g_1^{x_1} g_2^{x_2} \mod p$ *is a collision resistant hash function.*

Remarks:

- For a small example, let us pick $p = 5$ with primitive roots $g_1 = 2$ and $g_2 = 3$. We choose $x_1 = 3$ and $x_2 = 4$, obtaining the hash $h(3, 4) = 2^3 3^4 \mod 5 = 3 \mod 5$.
- Popular hash functions used in cryptography include the Secure Hash Algorithm (SHA) and the Message-Digest Algorithm (MD).
- It can be shown that finding a collision for the hash function described in Theorem 13.30 is equivalent to solving the discrete logarithm problem for $\log_{g_1} g_2$. Thus, the hash function is a collision resistant hash function, as we assume the discrete logarithm problem to be computationally hard.
- One might think that using a collision resistant hash function is good enough to store passwords for a service. E.g., store the hash of each password, and then compare it to the input of the user. Even if the hashes are leaked, an attacker Eve cannot recover the passwords – or can she?
- In practice, many users use short passwords, trading security for convenience. Eve can sample the hashes of common passwords such as "*password*", revealing the passwords of all users using these simple passwords. To counter this attack, one uses a technique called *salting*: The service adds a random bitstring (the *salt*) to each password before storing the hash (or, less secure, but simpler, the username). Even if the salt is known for each user, Eve needs to attack the hash of each user individually.
- To make life for Eve even harder, it is good practice to use hash functions that provably need a lot of computation and memory to execute. However, there is still a trade off as the real user wants to log in fast as well.
- Many web services already offer secure two-factor authentication (e.g., via mobile phones) instead of just passwords or challenge-response systems. However, there is a trade-off between security and convenience.
- Are we resistant against malleability now, if we include a hash of the encrypted message? No: An attacker changing the message can change the hash as well, as the hash function is not assumed to be secret. How do we prevent the hash from being modified without being noticed? The answer are HMACs:

Definition 13.31 (Message Authentication Code, MAC)**.** *A message authentication code is a bitstring that verifies that a message comes from the desired sender and was not changed until reaching the receiver.*

Definition 13.32 (Hash-Based Message Authentication Code, HMAC)**.** *A hash-based message authentication code is a MAC that uses a collision resistant hash function in combination with a secret key.*

Algorithm 13.33 Hash-Based Message Authentication Code Generation

Input: An encrypted message c, to be sent from Alice to Bob, the publicly known hash function h from Theorem 13.30, and a secret key $1 \leq k \leq c$ known to Alice and Bob.
Result: An HMAC for c, checkable by Bob.

1: Alice computes $h_A = h(k, h(k, c))$, and sends c, h_A to Bob
2: Bob computes $h_B = h(k, h(k, c))$, and checks if $h_A = h_B$

Remarks:

- In practice, if $k > c$, then k will be hashed to have a smaller size. Also, the key will be padded for extra security.
- If an attacker wants to change the message, he needs to change the HMAC too. To change the HMAC, he needs to know the secret key k
- Algorithm 13.33 can be also used with any other collision resistant hash function.

13.5 Public Key Cryptography

"*Love all, trust a few.*" – William Shakespeare

Definition 13.34 (Public Key Cryptography)**.** *A public key cryptography system uses two keys: A public key k_p, to be disseminated to everyone, and a secret (private) key k_s, only known to the owner. A message encrypted with the secret key can be decrypted with the corresponding public key. Analogously, a message encrypted with the public key can be decrypted with the corresponding secret key.*

Remarks:

- Popular public key cryptosystems include RSA and elliptic curve cryptography.
- With public key cryptography, we have reduced the number of keys – everyone just needs a secret and a public key.
- A conceptual way to think of public key cryptography is as follows: The secret key is a physical (secret) key that opens a specific type of padlock, and this type of padlock is freely available. The public key is a physical key too, freely available, but it opens only a (secret)

specific type of padlock. If Alice wants to send Bob an encrypted message, she applies his public padlock to the message container, and only Bob can open it. Similarly, if Alice wants to authenticate her message to Bob, she locks the container with her secret padlock, and only Alice's public key can unlock it. Lastly, if Alice wants to ensure both encryption and authentication, she applies both her own secret padlock and Bob's public padlock to the message container.

- We will now extend the Diffie-Hellman algorithm to public key cryptography, resulting in Algorithms 13.35 and 13.36.

- Recall the Diffie-Hellman Key Exchange algorithm (Algorithm 13.14). There, Alice picked a secret number k_A and computed a number A which she sent to Bob. We use the exact same idea in Algorithm 13.35 to generate a pair of a secret key and a public key which we call k_s and k_p, respectively.

Algorithm 13.35 Elgamal Key Generation

Input: Publicly known prime p and a primitive root g of p.
Result: Alice generates a public and a secret key

1: Alice randomly chooses $k_s \in \{1, 2, \ldots, p-1\}$ as her secret key
2: Alice calculates $k_p = g^{k_s} \mod p$ as her public key

Remarks:

- Alice can publish p, g, k_p, but must keep k_s to her own.

- We will explain public key encryption, before covering authentication.

- Recall how Bob calculated a key K from A and k_B in the Diffie-Hellman algorithm. In the context of public key encryption, when Bob wants to send an encrypted message to Alice, he calculates K from Alice's public key k_p and a secret random number x which he chooses himself.

- Bob then generates the number $c_1 = (g^x \mod p)$ and sends it to Alice, which she will later use to calculate K. Additionally, Bob sends c_2, which is his message m, encrypted using the key K.

- Once Alice has calculated K from c_1 and her secret key k_s, she can use K to reconstruct Bob's message m from c_2.

Theorem 13.37 (Fermat's little theorem)**.** *Let p be a prime number. Then, for any $a \in \mathbb{N}$ holds: $a^p = a \mod p$. If a is not divisible by p, then $a^{p-1} = 1 \mod p$.*

Theorem 13.38. *Algorithm 13.36 is correct.*

Algorithm 13.36 Elgamal Public Key Encryption and Decryption

Input: Alice and Bob know p, g, k_p, Alice knows k_s.
Result: Bob sends Alice an encrypted message, which she can decrypt.

1: Bob picks a message $m \in \{1, 2, \ldots, p-1\}$ and a random number $x \in \{1, 2, \ldots, p-1\}$ and calculates $K = (k_p)^x \mod p$.
2: Bob sends $c_1 = (g^x \mod p)$ and $c_2 = (K \cdot m \mod p)$ to Alice
3: Alice first calculates $K = (c_1)^{k_s} \mod p$
4: Alice then obtains $m = c_2 \cdot K^{p-2} \mod p$

Proof. We refer to the proof of Theorem 13.16 for the fact that Alice and Bob have calculated the same value for K.

It remains to show that Alice can restore the message m:

$$\begin{aligned} c_2 \cdot K^{p-2} &= (K \cdot m) \cdot K^{p-2} \mod p \\ &= m \cdot K^{p-1} \mod p \\ &= m \mod p \quad \text{(using Theorem 13.37)} \end{aligned}$$

□

Remarks:

- Note that the Elgamal encryption in Algorithm 13.36 on its own is malleable: An attacker can relay $c_1 = (g^x \mod p)$ as $z \cdot c_1$, resulting in a valid decryption of zm. In practice, techniques such as HMACs from the previous section should be used.
- Still, we can now send someone an encrypted message using public key cryptography, but what about authentication?
- Again, we first need some number theoretic preliminaries.

Definition 13.39 (Greatest Common Divisor, gcd)**.** *The greatest common divisor (gcd) of two integers i_1, i_2 is the largest integer that divides i_1 and i_2 without a remainder.*

Theorem 13.40. *Let p be a prime and i be an integer with $gcd(i, p) = 1$. Let $a_1, a_2 \in \mathbb{N}$. If $a_1 = a_2 \mod (p-1)$, then $i^{a_1} = i^{a_2} \mod p$.*

Remarks:

- A multiplicative inverse modulo p (in this algorithm: $x^{-1} \mod p$), can be calculated using, e.g., the extended Euclidean algorithm.

Lemma 13.42. *Algorithm 13.41 is correct.*

Proof. With $d = (m - ak_s)x^{-1} \mod (p-1)$, it follows that:

$$dx = m - ak_s \mod (p-1) \Rightarrow m = dx + ak_s \mod (p-1).$$

Using Theorem 13.40, we now obtain $g^{dx+ak_s} = g^m \mod p$. Hence,

$$k_p^a a^d \mod p = \left(g^{k_s}\right)^a (g^x)^d = g^{ak_s} g^{dx} \mod p = g^m \mod p.$$

□

Algorithm 13.41 Elgamal Authentication

Input: Alice and Bob know p, g, k_p, Alice knows k_s.
Result: Alice signs a message $m \in \{1, 2, \ldots, p-1\}$, which Bob authenticates.

1: Alice picks a random $x \in \{1, 2, \ldots, p-1\}$, with $gcd(x, p-1) = 1$
2: Alice calculates $a = g^x \mod p$ and $b = x^{-1} \mod (p-1)$
3: Alice calculates $d = (m - ak_s)b \mod (p-1)$
4: Alice sends the message m and the signature (a, d) to Bob
5: Bob checks if $1 \leq a \leq p-1$, else he rejects
6: Bob accepts Alice's signature for m if $k_p^a a^d = g^m \mod p$

Remarks:

- The security of the Elgamal public key cryptography again depends on the hardness of the discrete logarithm problem.
- We can now authenticate a message using public key cryptography, e.g., we can check that the public key of Alice corresponds to Alice's secret key.
- However, we are back still at our old problem: How do I know that Alice's public key really belongs to Alice? Maybe Eve pretended to be Alice? To use a famous saying by Peter Steiner: "*On the Internet, nobody knows you're a dog*".
- What can we do, unless we personally meet with everyone to exchange our public keys? The answer lies in trusting a few, in order to trust many: Let's say that you don't know Alice, but both Alice and you know Doris. If you trust Doris, then Doris can verify Alice's public key for you. In the future, you can ask Alice to vouch for her friends as well, etc.
- Trust is not limited to real persons though, especially since Alice and Doris are represented by their keys. Take a website like PayPal for example. How do you know that you give them your credit card information, and not some infamous Nigerian princess Eve? You probably don't know anybody who personally knows PayPal...

Definition 13.43 (Web of Trust)**.** *Let $G = (V, E)$ be a graph, where an edge between two nodes u, v represents trust between u, v. For any two nodes u, w, we say u trusts w if there is a path from u to w in G.*

Remarks:

- Hence, if you want someone to authenticate themselves, you need to find a path in the Web of Trust to them.
- In practice, the Web of Trust is a bit more sophisticated, as you can assign various levels of trust – and you might only trust someone in short distance.

- The whole situation is a bit of a chicken and egg dilemma though. In the beginning, you don't trust anyone, and nobody trusts you. You may want to find some well-connected nodes and gain their trust. This is the motivation for certificate authorities.

Definition 13.44 (Certificate Authority, CA)**.** *A certificate authority is a node in a web of trust that is trusted by many other nodes.*

Remarks:

- A main distinction between a CA (or nodes in general) and your real-life friends is that trust is not necessarily mutual, edges in the web of trust can also be directed. As such a node u might trust v, but v does not necessarily need to trust u.
- You will find trust for some certificate authorities pre-installed on your system/browser, known as root certificates. When you want to know if you can trust a node, the node can supply you with a path (chain of trust) from the CA. More specifically, you will be supplied with signatures which you can check (as you trust the CA).
- Again, one can implement various levels of trust, e.g., you might only trust short paths.
- Moreover, a CA might get compromised. This leads to the idea of key revocation, where one can check if a key for a signature has been compromised – the corresponding certificate can be generated by anyone holding the respective secret key. Another idea is to also generate expiration dates for keys.
- A totally different problem is that your own set of root certificates might be compromised, e.g., if malicious software adds new root certificates to one's device.

Chapter Notes

The concept of one-way functions is surprisingly old. In 1874, William Stanley Jevons wrote: "Can the reader say what two numbers multiplied together will produce the number 8616460799? I think it unlikely that anyone but myself will ever know." [9]. To spill the beans: $89681 \cdot 96079$.

The Diffie-Hellman Key Exchange was published in the seminal paper [6], parallel unpublished work also existed from Ellis et al. at the British intelligence service GCHQ. For some works showing the hardness of breaking the Diffie-Hellman key exchange, we refer to, e.g., [5], [11], [18]. For some more recommendations on how to choose the parameters of the Diffie-Hellman key exchange see RFC 3526 at `http://tools.ietf.org/html/rfc3526`. The currently fastest algorithms to solve the discrete logarithm problem still have non-practical runtime, e.g., [1]. The idea of challenging the other party to return an encrypted version of one's random number incremented by one in Algorithm 13.24 is taken from the Kerberos protocol. The Elgamal cryptosystem was published by Elgamal in 1984 [8], some years after RSA [15].

The first deterministic polynomial primality test, by Agrawal, Kayal, and Saxena, was published in [10], with an improved runtime of $\tilde{O}(\log^6 p)$ available at `https://math.dartmouth.edu/~carlp/aks041411.pdf`. The Miller-Rabin primality test is from Rabin [14] and Miller [12]. For an introduction to number theory, we recommend, e.g., [16].

The idea for the web of trust was proposed by Zimmermann in 1992. For certificate chains and key revocation, we refer to RFC 5280 at `http://tools.ietf.org/html/rfc5280`.

The Chaum-van-Heijst-Pfitzmann hash function described in Theorem 13.30 was published in [4] by Chaum et al., for the reduction to the discrete logarithm problem see, e.g., [19]. Although the runtime of the Chaum-van-Heijst-Pfitzmann hash function is too high in practice, it is chosen in this chapter as it is easier to understand compared to other related work. The subsequently described HMAC Algorithm 13.33 is from RFC 2104 at `https://tools.ietf.org/html/rfc2104`, with further security updates in RFC 6151, cf. `https://tools.ietf.org/html/rfc6151`.

The secret sharing variant discussed in this chapter is from Shamir [17], Blakley developed similar work in parallel in 1979 [3], and also discussed its relation to one-time pads [2].

While CBC seems superior to ECB, there is one downside: Decryption of ECB can be parallelized, but the decryption of CBC has to be sequential. The in this context mentioned AES encryption is a symmetric key algorithm, based on the Rijndael cipher of Daemen and Rijmen. Details of the Advanced Encryption Standard can be found in `http://csrc.nist.gov/publications/fips/fips197/fips-197.pdf`. AES, with a key length of 128, 192, or 256 bits, replaced DES (Data Encryption Standard), as its key length of just 56 was no longer secure enough against brute-force attacks.

For a general overview of the topic of computer security, we recommend [13] and [7]. Lastly, as a very general recommendation, we urge you not to implement your own cryptosystem unless you really know what you are doing – there is just too much that can easily be missed.

This chapter was written in collaboration with Klaus-Tycho Förster.

Bibliography

[1] Leonard Adleman. A subexponential algorithm for the discrete logarithm problem with applications to cryptography. In *Proceedings of the 20th Annual Symposium on Foundations of Computer Science*, SFCS '79, pages 55–60, Washington, DC, USA, 1979. IEEE Computer Society.

[2] G. R. Blakley. One time pads are key safeguarding schemes, not cryptosystems fast key safeguarding schemes (threshold schemes) exist. In *Proceedings of the 1980 IEEE Symposium on Security and Privacy, Oakland, California, USA, April 14-16, 1980*, pages 108–113. IEEE Computer Society, 1980.

[3] G.R. Blakley. Safeguarding cryptographic keys. In *Proceedings of the 1979 AFIPS National Computer Conference*, pages 313–317, Monval, NJ, USA, 1979. AFIPS Press.

[4] David Chaum, Eugène van Heijst, and Birgit Pfitzmann. Cryptographically strong undeniable signatures, unconditionally secure for the signer. In Joan Feigenbaum, editor, *Advances in Cryptology - CRYPTO '91, 11th Annual International Cryptology Conference, Santa Barbara, California, USA, August 11-15, 1991, Proceedings*, volume 576 of *Lecture Notes in Computer Science*, pages 470–484. Springer, 1991.

[5] Bert den Boer. Diffie-hillman is as strong as discrete log for certain primes. In Shafi Goldwasser, editor, *Advances in Cryptology - CRYPTO '88, 8th Annual International Cryptology Conference, Santa Barbara, California, USA, August 21-25, 1988, Proceedings*, volume 403 of *Lecture Notes in Computer Science*, pages 530–539. Springer, 1988.

[6] Whitfield Diffie and Martin E. Hellman. New directions in cryptography. *IEEE Trans. Information Theory*, 22(6):644–654, 1976.

[7] Niels Ferguson and Bruce Schneier. *Practical cryptography*. Wiley, 2003.

[8] Taher El Gamal. A public key cryptosystem and a signature scheme based on discrete logarithms. In G. R. Blakley and David Chaum, editors, *Advances in Cryptology, Proceedings of CRYPTO '84, Santa Barbara, California, USA, August 19-22, 1984, Proceedings*, volume 196 of *Lecture Notes in Computer Science*, pages 10–18. Springer, 1984.

[9] William Stanley Jevons. *The Principles of Science: A Treatise on Logic and Scientific Method*. Macmillan & Co., 1874.

[10] Nitin Saxena Manindra Agrawal, Neeraj Kayal. PRIMES Is in P. *Annals of Mathematics*, 160(2):781–793, 2004.

[11] Ueli M. Maurer. Towards the equivalence of breaking the diffie-hellman protocol and computing discrete algorithms. In Yvo Desmedt, editor, *Advances in Cryptology - CRYPTO '94, 14th Annual International Cryptology Conference, Santa Barbara, California, USA, August 21-25, 1994, Proceedings*, volume 839 of *Lecture Notes in Computer Science*, pages 271–281. Springer, 1994.

[12] Gary L. Miller. Riemann's hypothesis and tests for primality. *J. Comput. Syst. Sci.*, 13(3):300–317, December 1976.

[13] Josef Pieprzyk, Thomas Hardjono, and Jennifer Seberry. *Fundamentals of computer security*. Springer, 2003.

[14] M.O. Rabin. Probabilistic algorithms for testing primality. *J. Number Theory*, 12:128 – 138, 1980.

[15] R. L. Rivest, A. Shamir, and L. Adleman. A method for obtaining digital signatures and public-key cryptosystems. *Commun. ACM*, 21(2):120–126, February 1978.

[16] Winfried Scharlau and Hans Opolka. *From Fermat to Minkowski: lectures on the theory of numbers and its historical development*. Undergraduate Texts in Mathematics. Springer, New York, 1985.

[17] Adi Shamir. How to share a secret. *Commun. ACM*, 22(11):612–613, 1979.

[18] Victor Shoup. Lower bounds for discrete logarithms and related problems. In Walter Fumy, editor, *Advances in Cryptology - EUROCRYPT '97, International Conference on the Theory and Application of Cryptographic Techniques, Konstanz, Germany, May 11-15, 1997, Proceeding*, volume 1233 of *Lecture Notes in Computer Science*, pages 256–266. Springer, 1997.

[19] Douglas R. Stinson. *Cryptography - theory and practice.* Discrete mathematics and its applications series. CRC Press, 1995.