CS459/698 Privacy, Cryptography, Network and Data Security

Discrete Logarithm, Diffie-Hellman, ElGamal

 $h = g^x$, find x





I bet we can use that



But don't forget about me

Groups?

Groups - Sets with specific properties

A **group** is a set of elements (usually numbers) that are related to each other according to well-defined operations.

- Consider a multiplicative group Z_p^*
 - This boils down to the set of non-zero integers between 1 and p-1 modulo $p \rightarrow A$ finite group
 - For p = 5, we have group $Z_5^* = \{1,2,3,4\} \rightarrow i.e.$, the order **n** of Z_5^* is 4
 - In this group, operations are carried out mod 5:
 - 3 * 4 = 12 mod 5 = 2
 - $2^3 = 2 * 2 * 2 = 8 \mod 5 = 3$

Group axioms

To be a group, these sets should respect some axioms

- Closure
- Identity existence
- Associativity
- Inverse existence
- Groups can also be <u>commutative</u> and <u>cyclic</u> (up next)

Let's take a look at some of these axioms (using multiplication as the operation)

Closure

- For every x,y in the group, x * y is in the group
 - o i.e., the multiplication of two group elements falls within the group too

- Example:
 - o in Z_5^* , 2* 3 = 6 mod 5 = 1

Identity Existence

- There is an element e such that e * x = x * e = x
 - o i.e., has an element **e** such that any element times **e** outputs the element itself

- o In any Z_p^* , the identity element is 1
- \circ For $Z_5^* : 1 * 3 = 3 \mod 5 = 3$

Associativity

For any x, y, z in the group, (x * y) * z = x * (y * z)

o For
$$Z_5^*$$
: (2 * 3) * 4 = 1 * 4 = 2 * (3 * 4) = 2 * 2 = 4

Inverse Existence

For any x in the group, there is a y such that x * y = y * x = 1

- For $Z_5^*: 2*3 = 1, 3*2 = 1 (2 and 3 are inverses)$
- \circ 4 * 4 = 16 mod 5 = 1 (4 is its own inverse)

Abelian Groups

- Abelian groups are groups which are commutative
- This means that x * y = y * x for any group elements x and y

- Example:
 - \circ For $Z_5^*: 3*4=2, 4*3=2$

Cyclic groups

- A group is called **cyclic** if there is at least one element **g** such that its powers (g¹, g², g³, ...) mod p span all distinct group elements.
 - o **g** is called the "generator" of the group

- For Z_5^* , there are two generators (2 and 3):
 - $2^1 = 2$, $2^2 = 4$, $2^3 = 3$, $2^4 = 1$
 - \blacksquare 3¹ = 3, 3² = 4, 3³ = 2, 3⁴ = 1

Cyclic subgroups

We can have cyclic subgroups within larger finite groups

Example:

- The order of any cyclic subgroup of F_{607}^* must divide $\mathbf{n} = |F_{607}^*| = 606$
- \circ Thus, F_{607}^* has subgroups of orders {1, 2, 3, 6, 9, 18, 101, 202, 303, 606}

Important for later:

- The subgroup of order 101 is a subset of F_{607}^* . All calculations involving its generator **g** must take place in F_{607}^* , which uses modulo 607 arithmetic.
- Even though the subgroup has order n=101, its elements are still numbers in F_{607}^* , and their operations are also defined modulo 607.

 $h = g^x$, find x





I bet we can use that



But don't forget about me

$$h = g^x$$
, find x

Discrete: we are dealing with integers instead of real numbers

Logarithm: we are looking for the logarithm of **x** base **g**

$$\circ$$
 e.g., $\log_2 256 = 8$, since $2^8 = 256$

Given $(g,h) \in \mathbf{G} \times \mathbf{G}$, find $x \in \mathbf{Z_p}^*$ such that:

$$h = g^x$$

Here, **G** is a multiplicative group, just like we saw during the examples. (But **p** is thousands of bits long)

Solutions to the Discrete Logarithm Problem?

If there's one solution, there are infinitely many

(thank you Fermat's little theorem and modular arithmetic "wrap-around")

How to solve DLP in cyclic groups of prime order?

Is the group cyclic, finite, and abelian?

Has a generator that spans all elements

Has a limited number of elements

Multiplication is commutative



Baby-step/Giant-step algorithms!!!

Baby-Step/Giant-Step Algorithm?

- A cyclic group G = <g> which has prime order n
- $h \in G$, Goal: find $x \pmod{n}$ such that $h = g^x$
- Every element x ∈ G can be written as: x = i + j*[sqrt(n)]
 - For integers m, i, j satisfying $0 \le i, j \le m$.
 - \bigcirc m = [sqrt(n)]



Then:

$$h = g^{i + j*[sqrt(n)]}$$

$$g^{i} = h \cdot (g^{-[sqrt(n)]})^{j}$$

Baby-Step/Giant-Step Algorithm? Notation.

log_a x mod n is obtained by comparing two lists:

$$g^i = h \cdot (g^{-[sqrt(n)]})^j$$

When we find a coincidence, the equality holds and then x = i + j*[sqrt(n)]



Baby-step/Giant-Step Algorithm

1.
$$x = i + j*[sqrt(n)]$$



Baby-step/Giant-Step Algorithm

- 1. x = i + j*[sqrt(n)]
- 2. $0 \le i, j < [sqrt(n)]$



Baby-step/Giant-Step Algorithm

- 1. x = i + j*[sqrt(n)]
- 2. $0 \le i, j < [sqrt(n)]$



3. Baby-step: $g_i \leftarrow g^i$ for $0 \le i < [sqrt(n)]$

Baby-step/Giant-Step Algorithm

- 1. x = i + j*[sqrt(n)]
- 2. $0 \le i, j < [sqrt(n)]$

Produces pairs: (g_i,i)



3. Baby-step: $g_i \leftarrow g^i$ for $0 \le i < [sqrt(n)]$

Baby-step/Giant-Step Algorithm

- 1. x = i + j*[sqrt(n)]
- 2. $0 \le i, j < [sqrt(n)]$





4. Giant-step: $h_j \leftarrow h^*g^{-j \lceil sqrt(n) \rceil}$, for $0 \le j < \lceil sqrt(n) \rceil$



Baby-step/Giant-Step Algorithm

- 1. x = i + j*[sqrt(n)]
- 2. $0 \le i, j < [sqrt(n)]$





- 3. Baby-step: $g_i \leftarrow g^i$ for $0 \le i < [sqrt(n)]$
- 4. Giant-step: $h_i \leftarrow h^*g^{-j \lceil sqrt(n) \rceil}$, for $0 \le j < \lceil sqrt(n) \rceil$

Overall time and space O(Sqrt(n))

Baby-step/Giant-Step Algg

- 1. x = i + j*[sqrt(n)]
- 2. $0 \le i, j < [sqrt(n)]$

4. Gi2

Note: For DLP in group G to be "difficult enough" (e.g., 2128 order order operations), needs prime order operations), subgroup of size greater than 2256

verall time a

space O(Sqrt(n))

 $(11_j,j)$

• Consider the subgroup of prime order 101 ($\mathbf{n} = 101$) in F_{607}^* , generated by $\mathbf{g} = 64$

i	64 ⁱ (mod 607)	i	""
0		6	
1		7	
2		8	
3		9	
4		10	
5		-	

Take that we know this...



Focusing on the subgroup **ensures** that every element in the problem is generated by the **known** g=64, making it possible to **solve** the DLP.

• Consider the subgroup of prime order 101 ($\mathbf{n} = 101$) in F_{607}^* , generated by $\mathbf{g} = 64$

i	64 ⁱ (mod 607)	i	66 33
0		6	
1		7	
2		8	
3		9	
4		10	
5		-	

Take that we know this...

Focusing on the subgroup **ensures** that every element in the problem is generated by the **known** g=64, making it possible to **solve** the DLP.



This tells us x is in the range $0 \le x < 101$ because the subgroup has order 101.

• Consider the subgroup of prime order 101 ($\mathbf{n} = 101$) in F_{607}^* , generated by $\mathbf{g} = 64$

i	64 ⁱ (mod 607)	i	""
0		6	
1		7	
2		8	
3		9	
4		10	
5		_	

Take that we know this...

Focusing on the subgroup **ensures** that every element in the problem is generated by the **known** g=64, making it possible to **solve** the DLP.

This tells us x is in the range $0 \le x < 101$ because the subgroup has order 101.



But recall we're operating in mod 607 due to F_{607}^*

• Consider the subgroup of prime order 101 ($\mathbf{n} = 101$) in F_{607}^* , generated by $\mathbf{g} = 64$

i	64 ⁱ (mod 607)	i	66 33
0		6	
1		7	
2		8	
3		9	
4		10	
5		_	



Baby-step: $g_i \leftarrow g^i$ for $0 \le i < [sqrt(n)]$

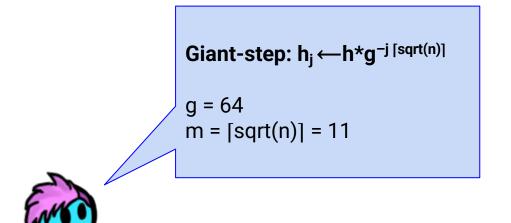
i	64 ⁱ (mod 607)	i	""
0	1	6	330
1	64	7	482
2	454	8	498
3	527	9	308
4	343	10	288
5	100	-	



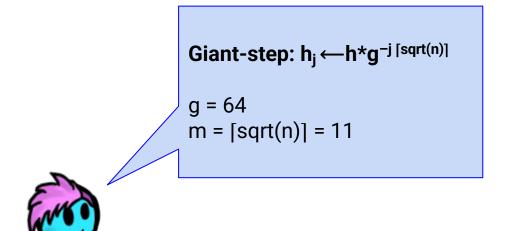
Baby-step:
$$g_i \leftarrow g^i$$
 for $0 \le i < \lceil sqrt(n) \rceil$

$$g = 64$$

m = [sqrt(n)] = 11

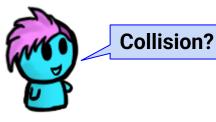


i	182* 64 ^{-11*j} (mod 607)	i	
0		6	
1		7	
2		8	
3		9	
4		10	
5		-	



i	182* 64 ^{-11*j} (mod 607)	i	
0	182	6	60
1	143	7	394
2	69	8	483
3	271	9	76
4	343	10	580
5	573	-	

i		i	64 ⁱ (mod 607)
0	1	6	330
1	64	7	482
2	454	8	498
3	527	9	308
4	343	10	288
5	100	-	



j		j	182* 64 ^{-11*j} (mod 607)
0	182	6	60
1	143	7	394
2	69	8	483
3	271	9	76
4	343	10	580
5	573	-	

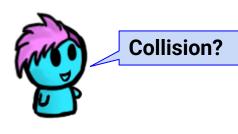
i		i	64 ⁱ (mod 607)
0	1	6	330
1	64	7	482
2	454	8	498
3	527	9	308
4	343	10	288
5	100	-	



j		j	182* 64 ^{-11*j} (mod 607)
0	182	6	60
1	143	7	394
2	69	8	483
3	271	9	76
4	343	10	580
5	573	-	

DLP Example, $182 = 64^{x} \pmod{607}$

i		i	64 ⁱ (mod 607)
0	1	6	330
1	64	7	482
2	454	8	498
3	527	9	308
4	343	10	288



j		j	182* 64 ^{-11*j} (mod 607)
0	182	6	60
1	143	7	394
2	69	8	483
3	271	9	76
4	343	10	580

Match when **i=4** and **j=4**.

(i is not necessarily equal to j, but it happened on this run $^-_(^\vee)_-/^-$

DLP Example, $182 = 64^{x} \pmod{607}$

i		i	64 ⁱ (mod 607)
0	1	6	330
1	64	7	482
2	454	8	498
3	527	9	308
4	343	10	288

5

100

x = i + j*[sqrt(n)]



Collision?

j		j	182* 64 ^{-11*j} (mod 607)
0	182	6	60
1	143	7	394
2	69	8	483
3	271	9	76
4	343	10	580

Recall: x = i + j*[sqrt(n)]

So: x = 4 + 4*11 = 48.

DLP Example, $182 = 64^{x} \pmod{607}$

i		i	64 ⁱ (mod 607)
0	1	6	330
1	64	7	482
2	454	8	498
3	527	9	308
4	343	10	288

5

100



Collision?

j		j	182* 64 ^{-11*j} (mod 607)
0	182	6	60
1	143	7	394
2	69	8	483
3	271	9	76

Verify: 64⁴⁸ (mod 607) = 182

Recall: x = i + j*[sqrt(n)]

So: x = 4 + 4*11 = 48.

Diffie-Hellman



A public-key protocol published in 1976 by Whitfield Diffie and Martin Hellman



Allows two parties that have no prior knowledge of each other to jointly establish a shared secret key over an insecure channel



Key used to encrypt subsequent communications using a symmetric key cipher

- Used for establishing a <u>shared secret</u> (lacks authentication; we'll see why this is <u>bad</u>)
- Assume as public parameters generator g and prime p
- Alice (resp. Bob) generates private value a (resp. b)

- Used for establishing a <u>shared secret</u> (lacks authentication; we'll see why this is <u>bad</u>)
- Assume as public parameters generator g and prime p
- Alice (resp. Bob) generates private value a (resp. b)



$$B^{a} = (g^{b})^{a} = g^{ba}$$
 $A^{b} = (g^{a})^{b} = g^{ab}$

Alice and Bob can derive the same value by exchanging public values and combining them with their private ones!

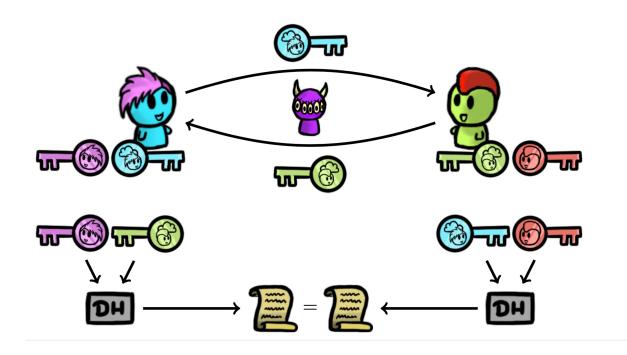
- Used for establishing a <u>shared secret</u> (lacks authentication; we'll see why this is <u>bad</u>)
- Assume as public parameters generator g and prime p
- Alice (resp. Bob) generates private value a (resp. b)



$$B^{a} = (g^{b})^{a} = g^{ba}$$
 $A^{b} = (g^{a})^{b} = g^{ab}$

Resist keying temptation: the shared value should not <u>immediately</u> be used as a key. Gab is a random element inside a group, but not necessarily a random bit string

<u>Diffie-Hellman Key Exchange – Visualization</u>



Diffie-Hellman relies on the DLP

DH can be broken by recovering the private value **a** from the public value **g**^a

The adversary must not be able to solve the DLP



The Decisional Diffie-Hellman Problem

Given **g**, **g**^a, **g**^b distinguish **g**^{ab} from random **g**^c

- An adversary should NOT be able to learn anything about the secret g^{ab} after observing public values g^a and g^b
 - Assume g^{ab} and g^c occur with the same probability

The Decisional Diffie-Hellman Problem

Given **g**, **g**^a, **g**^b distinguish **g**^{ab} from random **g**^c

- An adversary should NOT be able to learn anything about the secret g^{ab} after observing public values g^a and g^b
 - Assume g^{ab} and g^c occur with the same probability

Useful assumption **beyond** DH key exchange!



EIGamal relies on the DDH assumption

ElGamal

ElGamal Public Key Cryptosystem

- Let p be a prime such that the DLP in $(\mathbf{Z}_{p}^{*,})$ is infeasible
- Let α be a generator in \mathbf{Z}_{p}^{*} and \mathbf{a} a secret value
- **PubK** ={ (p,α,β) : $\beta \equiv \alpha^a \pmod{p}$ }

- For message **m** and secret random **k** in Z_{p-1} :
 - \circ e_K(m,k) = (y₁, y₂), where $\mathbf{y_1} = \alpha^k \mod p$ and $\mathbf{y_2} = m\beta^k \mod p$

- For y_1 , y_2 in \mathbf{Z}_p^* :
 - \bigcirc d_K(y₁, y₂)= y₂(y₁^a)⁻¹ mod p

ElGamal: The Keys

- 1. Bob picks a "large" prime \mathbf{p} and a generator $\mathbf{\alpha}$.
 - a. Assume message m is an integer 0 < m < p
- 2. Bob picks secret integer a
- 3. Bob computes $\beta \equiv \alpha^a \pmod{p}$



ElGamal: The Keys

- 1. Bob picks a "large" prime \mathbf{p} and a generator $\mathbf{\alpha}$.
 - a. Assume message m is an integer 0 < m < p
- 2. Bob picks secret integer a
- 3. Bob computes $\beta \equiv \alpha^a \pmod{p}$
- 4. Bob's public key is (p, α, β)





ElGamal: The Keys

- 1. Bob picks a "large" prime \mathbf{p} and a generator $\mathbf{\alpha}$.
 - a. Assume message m is an integer 0 < m < p
- 2. Bob picks secret integer a
- 3. Bob computes $\beta \equiv \alpha^a \pmod{p}$
- 4. Bob's public key is (**p**, **α**, **β**) **6**









ElGamal: Encryption

Bob's $Pub_K \rightarrow (p, \alpha, \beta)$ Bob's $Priv_K \rightarrow a$ $\beta \equiv \alpha^a \pmod{p}$





I choose secret integer **k**

ElGamal: Encryption

Bob's $Pub_K \rightarrow (p, \alpha, \beta)$ Bob's $Priv_K \rightarrow a$ $\beta \equiv \alpha^a \pmod{p}$





I choose secret integer **k**

Compute $\mathbf{y}_1 \equiv \alpha^k \pmod{p}$

ElGamal: Encryption

Bob's $Pub_K \rightarrow (p, \alpha, \beta)$ Bob's $Priv_K \rightarrow a$ $\beta \equiv \alpha^a \pmod{p}$





I choose secret integer **k**

Compute $\mathbf{y}_1 \equiv \alpha^k \pmod{p}$

Compute $\mathbf{y_2} \equiv \beta^k \text{ m (mod p)}$

Bob's $Pub_K \rightarrow (p, \alpha, \beta)$

Bob's $Priv_K \rightarrow a$

 $\beta \equiv \alpha^a \pmod{p}$



ElGamal: Encryption

I choose secret integer k



Compute $\mathbf{y}_1 \equiv \alpha^k \pmod{p}$

Compute $\mathbf{y}_2 \equiv \beta^k \text{ m (mod p)}$

Send y_1 and y_2 to Bob





ElGamal: Decryption

I choose secret integer k



Compute $\mathbf{y}_1 \equiv \alpha^k \pmod{p}$

Compute $\mathbf{y_2} \equiv \beta^k \text{ m (mod p)}$

Send y_1 and y_2 to Bob

Compute $\mathbf{y_1y_2}^{-a} \equiv m \pmod{p}$





ElGamal: Decryption

I choose secret integer k



Compute $\mathbf{y}_1 \equiv \alpha^k \pmod{p}$

Compute $\mathbf{y_2} \equiv \beta^k \text{ m (mod p)}$

Send y_1 and y_2 to Bob

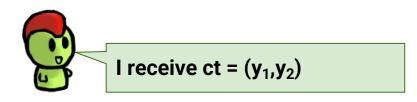
Compute $y_1y_2^{-a} \equiv m \pmod{p}$



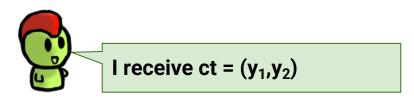
Bob can <u>decrypt</u> since:

$$\mathbf{y_2y_1}^{-a} \equiv \beta^k \ m \ (\alpha^k)^{-a} \equiv m \ (mod \ p)$$

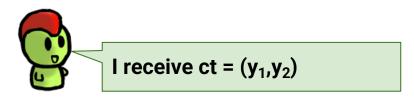
• The plaintext m is "hidden" by multiplying it by β^k to get y_2



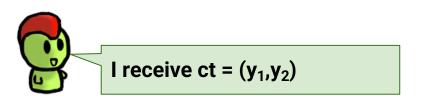
- The plaintext m is "hidden" by multiplying it by β^k to get y_2
- The ciphertext includes α^k so that Bob can compute β^k from α^k (because Bob knows a)



- The plaintext m is "hidden" by multiplying it by β^k to get y₂
- The ciphertext includes α^k so that Bob can compute β^k from α^k (because Bob knows a)
- Thus, Bob can "reveal" m by dividing y_2 by β^k



- The plaintext m is "hidden" by multiplying it by β^k to get y_2
- The ciphertext includes α^k so that Bob can compute β^k from α^k (because Bob knows a)
- Thus, Bob can "reveal" m by dividing y₂ by β^k





Let's see an example!

Example

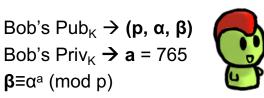
• Let p=2579, $\alpha = 2$, $\beta = 2^{765} \mod 2579 = 949$

Example

• Let p=2579, $\alpha = 2$, $\beta = 2^{765} \mod 2579 = 949$



I want to send **m**=1299 to Bob. I choose **k** = 853 for my random integer



Example

• Let p=2579, $\alpha = 2$, $\beta = 2^{765} \mod 2579 = 949$



I want to send m=1299 to Bob. I choose k = 853 for my random integer

$$\mathbf{y}_1 \equiv \mathbf{a}^{\mathbf{k}} \pmod{\mathbf{p}}$$

$$\mathbf{y_2} \equiv \beta^k \text{ m (mod p)}$$

Example

• Let p=2579, $\alpha = 2$, $\beta = 2^{765} \mod 2579 = 949$



I want to send **m**=1299 to Bob. I choose **k** = 853 for my random integer

$$\mathbf{y_1} \equiv \mathbf{\alpha^k} \pmod{\mathbf{p}}$$

$$\mathbf{y_2} \equiv \beta^k \text{ m (mod p)}$$

•
$$\mathbf{y_1} = 2^{853} \mod 2579 = 435$$

• $y_2 = 1299 * 949^{853} \mod 2579 = 2396$

Send y₁, y₂ to Bob

Example

- Bob now has y₁ and y₂
 - $y_1 = 2^{853} \mod 2579 = 435$
 - $y_2 = 1299 * 949^{853} \mod 2579 = 2396$



I received y = (435, 2396)

Example

- Bob now has y_1 and y_2
 - $y_1 = 2^{853} \mod 2579 = 435$
 - $y_2 = 1299 * 949^{853} \mod 2579 = 2396$



I received y = (435, 2396)

 $\mathbf{y_2y_1}^{-\mathbf{a}} \equiv \beta^k \ \mathbf{m} \ (\alpha^k)^{-\mathbf{a}} \equiv \mathbf{m} \ (\mathbf{mod} \ \mathbf{p})$

• m = 2396 * 435⁻⁷⁶⁵ mod 2759 = 1299



- Bob now has y₁ and y₂
 - $y_1 = 2^{853} \mod 2579 = 435$
 - $y_2 = 1299 * 949^{853} \mod 2579 = 2396$



I received y = (435, 2396)

 $\mathbf{y_2y_1}^{-a} \equiv \beta^k \ m \ (\alpha^k)^{-a} \equiv m \ (mod \ p)$

m = 2396 * 435⁻⁷⁶⁵ mod 2759 = 1299



Nice! That's the plaintext I wanted to send.



Example

- Bob now has y₁ and y₂
 - $y_1 = 2^{853} \mod 2579 = 435$
 - $y_2 = 1299 * 949^{853} \mod 2579 = 2396$



I received y = (435, 2396)

 $\mathbf{y_2y_1}^{-a} \equiv \beta^k \ m \ (\alpha^k)^{-a} \equiv m \ (mod \ p)$

• m = 2396 * 435⁻⁷⁶⁵ mod 2759 = 1299



Nice! That's the plaintext I wanted to send.



Insecure if the adversary can compute $\mathbf{a} = \log_{\alpha} \beta$



Example

- Bob now has y₁ and y₂
 - $y_1 = 2^{853} \mod 2579 = 435$
 - $y_2 = 1299 * 949^{853} \mod 2579 = 2396$



I received y = (435, 2396)

 $\mathbf{y_2y_1}^{-\mathbf{a}} \equiv \beta^k \ m \ (\alpha^k)^{-\mathbf{a}} \equiv m \ (mod \ p)$

• m = 2396 * 435⁻⁷⁶⁵ mod 2759 = 1299



Nice! That's the plaintext I wanted to send.



Insecure if the adversary can compute $\mathbf{a} = \log_{\alpha} \beta$

To be secure, DLP must be infeasible in \mathbf{Z}_p^*

But... We had RSA, why do we need ElGamal?

Extensions

- ElGamal supports Elliptic Curve Cryptography (ECC)
- Stronger security with smaller keys compared to RSA

Probabilistic Encryption

Adds semantic security with randomization (different ciphertexts for the same plaintext).

Homomorphic properties

o Additive homomorphism vs. RSA's multiplicative homomorphism

Network Security - Next class