# CMSC/Math 456: <br> Cryptography (Fall 2023) <br> Lecture 16 <br> Daniel Gottesman 

## Administrative

I am not sure if there will be a problem set this week.
Next week I will be away, and there will be guest lectures by Gorjan Alagic, topic TBA.

## Public Key Encryption



Public-key encryption is an asymmetric protocol.

## Public Key Encryption



Public-key encryption is an asymmetric protocol.
This class is being recorded

## Public Key Encryption




Bob


Eve
Public-key encryption is an asymmetric protocol.

## Public Key Encryption



Eve
Public-key encryption is an asymmetric protocol.
This class is being recorded

## Public Key Encryption



Eve
Public-key encryption is an asymmetric protocol.
This class is being recorded

## Public Key Encryption



Public-key encryption is an asymmetric protocol.
This class is being recorded

## Public Key Encryption



Public-key encryption is an asymmetric protocol.
This class is being recorded

## Definition of Public Key Encryption

Definition:A public-key encryption protocol is a set of three probabilistic polynomial-time algorithms (Gen, Enc, Dec).

Gen is the key generation algorithm. It takes as input s, the security parameter, and outputs a public key, private key pair $(e, d) \in\{0,1\}^{*} \times\{0,1\}^{*}$.

Enc is the encryption algorithm. It takes as input e and a plaintext or message $m \in\{0,1\}^{*}$ and outputs a ciphertext $c \in\{0,1\}^{*}$.

Dec is the decryption algorithm. It takes as input d and c and outputs some $m^{\prime} \in\{0,1\}^{*}$.

The encryption protocol is correct if

$$
\operatorname{Dec}(d, \operatorname{Enc}(e, m))=m
$$

Note: Gen here is much more complex than for private-key encryption and doesn't just generate random bit strings.

This class is being recorded

## Security of Public Key Encryption

Recall that EAV security for public key encryption is the same as CPA security and that both are defined by a game similar to the private key case:


The main difference from private keys is that now Eve is given the public key e. She can use e to choose the messages she wants to use as challenges as well as in her attack to decipher the message.

This class is being recorded

## Public Key Security

Definition: A public-key encryption protocol (Gen, Enc, Dec) with security parameter s has is EAV-secure and CPA-secure if, for any pair of messages $m_{0}$ and $m_{1}$ chosen by the adversary (using efficient algorithm $\mathscr{B}(e, s)$ ) and for any efficient attack $\mathscr{A}(e, c)$,
$\left|\operatorname{Pr}_{(e, d)}\left(\mathscr{A}\left(e, \operatorname{Enc}\left(e, m_{0}\right)\right)=1\right)-\operatorname{Pr}_{(e, d)}\left(\mathscr{A}\left(e, \operatorname{Enc}\left(e, m_{1}\right)\right)=1\right)\right| \leq \epsilon(s)$
for negligible $\epsilon(s)$ and probability taken over valid public key, private key pairs (e,d) and randomness of Enc, $\mathscr{A}$, and $\mathscr{B}$.


This class is being recorded

## RSA Encryption (Picture)

In RSA, Gen creates a public key ( $\mathrm{N}, \mathrm{e}$ ) and a private key ( $\mathrm{N}, \mathrm{d}$ ) with $e d=1 \bmod \varphi(N)$.


Bob
"RSA" stands for the inventors: Rivest, Shamir, and Adleman

## RSA Encryption (Picture)

In RSA, Gen creates a public key ( $\mathrm{N}, \mathrm{e}$ ) and a private key ( $\mathrm{N}, \mathrm{d}$ ) with $e d=1 \bmod \varphi(N)$.


Bob
"RSA" stands for the inventors: Rivest, Shamir, and Adleman

## RSA Encryption (Picture)

In RSA, Gen creates a public key ( $\mathrm{N}, \mathrm{e}$ ) and a private key ( $\mathrm{N}, \mathrm{d}$ ) with $e d=1 \bmod \varphi(N)$.


Bob
"RSA" stands for the inventors: Rivest, Shamir, and Adleman

## RSA Encryption (Picture)

In RSA, Gen creates a public key ( $\mathrm{N}, \mathrm{e}$ ) and a private key ( $\mathrm{N}, \mathrm{d}$ ) with $e d=1 \bmod \varphi(N)$.

$(\mathrm{N}, \mathrm{e})$

"RSA" stands for the inventors: Rivest, Shamir, and Adleman

## RSA Encryption (Gen, Enc, Dec)

Gen: Generate two random primes $p$ and $q$ which are $s$ bits long. Let $\mathrm{N}=\mathrm{pq}$. Choose $e, d \in \mathbb{Z}_{\varphi(N)}^{*}$ such that $e d=1 \bmod \varphi(N)$. The public key is $(\mathrm{N}, \mathrm{e})$ and the private key is ( $\mathrm{N}, \mathrm{d}$ ).
Enc: Given message $m$ and public key ( $\mathrm{N}, \mathrm{e}$ ). The ciphertext is $c=m^{e} \bmod N$.

Dec: Given ciphertext c and private key ( $\mathrm{N}, \mathrm{d}$ ). The decrypted message is $m^{\prime}=c^{d} \bmod N$.

## RSA Encryption (Gen, Enc, Dec)

Gen: Generate two random primes p and q which are s bits long. Let $\mathrm{N}=\mathrm{pq}$. Choose $e, d \in \mathbb{Z}_{\varphi(N)}^{*}$ such that $e d=1 \bmod \varphi(N)$. The public key is $(\mathrm{N}, \mathrm{e})$ and the private key is ( $\mathrm{N}, \mathrm{d}$ ).
Enc: Given message $m$ and public key ( $\mathrm{N}, \mathrm{e}$ ). The ciphertext is $c=m^{e} \bmod N$.

Dec: Given ciphertext c and private key ( $\mathrm{N}, \mathrm{d}$ ). The decrypted message is $m^{\prime}=c^{d} \bmod N$.

## Example:

$$
\begin{aligned}
& \text { Gen: Let } \mathrm{p}=\mathrm{I} \mathrm{I}, \mathrm{q}=17 \text {. Then } \mathrm{N}=187 \text { and } \\
& \varphi(N)=(p-1)(q-1)=10 \cdot 16=160 \text {. Let } \mathrm{e}=3 \text {; then } \\
& \mathrm{d}=107(\text { so } 3 \cdot 107=1 \bmod 160) .
\end{aligned}
$$

$$
\text { Enc: Let } \mathrm{m}=113 \text {. Then } c=113^{3} \bmod 187=5 .
$$

$$
\text { Dec: } m^{\prime}=5^{107} \bmod 187=113
$$

## Why is RSA Correct?

Why does $\operatorname{Dec}(d, \operatorname{Enc}(e, m))=m$ ?
Recall the Euler-Fermat theorem: $x^{\varphi(N)}=1 \bmod N$.
Then

$$
\operatorname{Dec}(d, \operatorname{Enc}(e, m))=m^{e d} \bmod N
$$

But

$$
e d=1+k \varphi(N)
$$

so

$$
\begin{aligned}
\operatorname{Dec}(d, \operatorname{Enc}(e, m)) & =m^{1+k \varphi(N)} \bmod N \\
& =m \cdot\left(m^{\varphi(N)}\right)^{k} \bmod N \\
& =m \cdot 1^{k} \bmod N
\end{aligned}
$$

## Picking e and d

Gen: Generate two random primes p and q which are s bits long. Let $\mathrm{N}=\mathrm{pq}$. Choose $e, d \in \mathbb{Z}_{\varphi(N)}^{*}$ such that $e d=1 \bmod \varphi(N)$. The public key is $(\mathrm{N}, \mathrm{e})$ and the private key is $(\mathrm{N}, \mathrm{d})$.

How do we pick e and d?

## Picking e and d

Gen: Generate two random primes p and q which are s bits long. Let $\mathrm{N}=\mathrm{pq}$. Choose $e, d \in \mathbb{Z}_{\varphi(N)}^{*}$ such that $e d=1 \bmod \varphi(N)$. The public key is $(\mathrm{N}, \mathrm{e})$ and the private key is $(\mathrm{N}, \mathrm{d})$.

How do we pick e and d?
I. From p and q , we can find $\varphi(N)=(p-1)(q-1)$.

## Picking e and d

Gen: Generate two random primes p and q which are s bits long. Let $\mathrm{N}=\mathrm{pq}$. Choose $e, d \in \mathbb{Z}_{\varphi(N)}^{*}$ such that $e d=1 \bmod \varphi(N)$.
The public key is $(\mathrm{N}, \mathrm{e})$ and the private key is $(\mathrm{N}, \mathrm{d})$.
How do we pick e and d?
I. From p and q , we can find $\varphi(N)=(p-1)(q-1)$.
2. Choose a convenient value for $e \in \mathbb{Z}_{\varphi(N)}^{*}$. It doesn't need to be random. A common choice is $\mathrm{e}=3$.

## Picking e and d

Gen: Generate two random primes p and q which are s bits long. Let $\mathrm{N}=\mathrm{pq}$. Choose $e, d \in \mathbb{Z}_{\varphi(N)}^{*}$ such that $e d=1 \bmod \varphi(N)$.
The public key is $(\mathrm{N}, \mathrm{e})$ and the private key is $(\mathrm{N}, \mathrm{d})$.
How do we pick e and d?
I. From p and q , we can find $\varphi(N)=(p-1)(q-1)$.
2. Choose a convenient value for $e \in \mathbb{Z}_{\varphi(N)}^{*}$. It doesn't need to be random. A common choice is $\mathrm{e}=3$.
3. Then use Euclid's algorithm. Since $e \in \mathbb{Z}_{\varphi(N)}^{*}$, e and $\varphi(N)$ are relatively prime, so we find $X, Y$ such that:

$$
X e+Y \varphi(N)=1
$$

## Picking e and d

Gen: Generate two random primes $p$ and $q$ which are $s$ bits long. Let $\mathrm{N}=\mathrm{pq}$. Choose $e, d \in \mathbb{Z}_{\varphi(N)}^{*}$ such that $e d=1 \bmod \varphi(N)$.
The public key is $(\mathrm{N}, \mathrm{e})$ and the private key is $(\mathrm{N}, \mathrm{d})$.
How do we pick e and d?
I. From p and q , we can find $\varphi(N)=(p-1)(q-1)$.
2. Choose a convenient value for $e \in \mathbb{Z}_{\varphi(N)}^{*}$. It doesn't need to be random. A common choice is $\mathrm{e}=3$.
3. Then use Euclid's algorithm. Since $e \in \mathbb{Z}_{\varphi(N)}^{*}$, e and $\varphi(N)$ are relatively prime, so we find $X, Y$ such that:

$$
X e+Y \varphi(N)=1
$$

4. Then we can let $\mathrm{d}=\mathrm{X}$, because

$$
X e=1-Y \varphi(N)=1 \bmod \varphi(N)
$$

## RSA 0\&A

- Why pick p and q instead of picking N directly?


## RSA 0\&A

- Why pick p and q instead of picking N directly?

To pick e and d, we need to know $\varphi(N)$, and to find $\varphi(N)$, we need to know the prime factorization of $N$. Factoring seems hard, so we pick the prime factors directly.

## RSA O\&A

- Why pick $p$ and $q$ instead of picking $N$ directly?

To pick e and d, we need to know $\varphi(N)$, and to find $\varphi(N)$, we need to know the prime factorization of $N$. Factoring seems hard, so we pick the prime factors directly.

- Eve knows e. Can't she also apply Euclid's algorithm to find e?


## RSA O\&A

- Why pick p and q instead of picking N directly?

To pick e and d, we need to know $\varphi(N)$, and to find $\varphi(N)$, we need to know the prime factorization of $N$. Factoring seems hard, so we pick the prime factors directly.

- Eve knows e. Can't she also apply Euclid's algorithm to find e?

Only if she knows $\varphi(N)$. But to compute $\varphi(N)$, it seems she needs to know the prime factorization of N . And factoring seems hard.

## Hardness of Factoring

Factoring a product of two large primes seems to be hard.
Definition: Given a security parameter $s$, let $P(N)$ be a probability distribution over (2s)-bit numbers. We say that factoring is average-case hard for $\mathrm{P}(\mathrm{N})$ if for any polynomial time algorithm $\mathscr{A}$, for random $N$ chosen according to $P(N)$,

$$
\operatorname{Pr}(\mathscr{A}(N) \text { succeeds }) \leq \epsilon(s)
$$

for $\epsilon(s)$ a negligible function, where we say $\mathscr{A}(N)$ succeeds if $\mathscr{A}(N)=p$ with p a non-trivial factor of $\mathrm{N}: 1<p<N$ and $p \mid N$.

When $N$ is chosen to be the product of two random s-bit primes, it appears that factoring is hard for the resulting distribution.

## Factoring and RSA

If Eve can factor N , she can break RSA:
Eve factors $\mathrm{N}=\mathrm{pq}$ and computes $\varphi(N)=(p-1)(q-1)$. Using Euclid's algorithm, she computes dgiven e and $\varphi(N)$.

## Factoring and RSA

If Eve can factor $N$, she can break RSA:
Eve factors $\mathrm{N}=\mathrm{pq}$ and computes $\varphi(N)=(p-1)(q-1)$. Using Euclid's algorithm, she computes d given e and $\varphi(N)$.

Theorem: Given N, if Eve knows $d$ and e, she has an efficient probabilistic algorithm to factor N .

Factoring N


Learning the private key for RSA

## Factoring and RSA

If Eve can factor N , she can break RSA:
Eve factors $\mathrm{N}=\mathrm{pq}$ and computes $\varphi(N)=(p-1)(q-1)$. Using Euclid's algorithm, she computes dgiven e and $\varphi(N)$.

Theorem: Given N, if Eve knows d and e, she has an efficient probabilistic algorithm to factor N .

Factoring N


Learning the private key for RSA

But perhaps it is possible to break RSA without learning the private key ....

## Is RSA Secure?

Vote: Is RSA as l've presented it CPA-secure? (Yes/No/Unknown)

## Is RSA Secure?

Vote: Is RSA as l've presented it CPA-secure? (Yes/No/Unknown)

Answer: No.
This version of RSA (Plain RSA) is deterministic: It cannot be CPA-secure.

If Eve is trying to determine which of two messages was sent, she can just try encrypting each of them using the public key. In Plain RSA, the message $m_{i}$ always corresponds to ciphertext $c_{i}=m_{i}^{e} \bmod N$, so she can easily find $i$ by matching $c=c_{i}$.

## Is RSA Secure?

Vote: Is RSA as l've presented it CPA-secure? (Yes/No/Unknown)

Answer: No.
This version of RSA (Plain RSA) is deterministic: It cannot be CPA-secure.

If Eve is trying to determine which of two messages was sent, she can just try encrypting each of them using the public key. In Plain RSA, the message $m_{i}$ always corresponds to ciphertext $c_{i}=m_{i}^{e} \bmod N$, so she can easily find $i$ by matching $c=c_{i}$.

We need to add a random element to Enc.

## Padded RSA

One option to add randomness to RSA is to pad the message with random bits:

Gen remains the same.
Enc: Given message m , generate $\ell$-bit random r and let $\tilde{m}=r \| m$ (the concatenation $r$ followed by $m$ ). Then the ciphertext $c=\tilde{m}^{e} \bmod N$.
Dec: Given c, compute $\tilde{m}^{\prime}=c^{d} \bmod N$. Discard the first $\ell$ bits of $\tilde{m}^{\prime}$ to get $m^{\prime}$.

## Padded RSA

One option to add randomness to RSA is to pad the message with random bits:

Gen remains the same.
Enc: Given message m , generate $\ell$-bit random r and let $\tilde{m}=r \| m$ (the concatenation r followed by m ). Then the ciphertext $c=\tilde{m}^{e} \bmod N$.
Dec: Given c, compute $\tilde{m}^{\prime}=c^{d} \bmod N$. Discard the first $\ell$ bits of $\tilde{m}^{\prime}$ to get $\mathrm{m}^{\prime}$.

## Example:

$$
\begin{aligned}
& \text { Gen: } \mathrm{N}=187, \mathrm{e}=3 \text {, and } \mathrm{d}=107 \text { as before. } \\
& \text { Enc: Let } m=15=1111_{2} \text {. Pad using } r=110_{2} \text {. Then } \\
& \tilde{m}=1101111_{2}=111 \text { and } c=111^{3} \bmod 187=100 \text {. } \\
& \text { Dec: } \tilde{m}^{\prime}=100^{107} \bmod 187=111 \text { and } m^{\prime}=15 \text {. }
\end{aligned}
$$

## Security of Padded RSA

If $\ell=\left\lfloor\log _{2} N\right\rfloor-1$, Padded RSA is CPA-secure assuming security of an RSA assumption: It is hard to find $x$ such that $x^{e}=y \bmod N$.

Note:This is a stronger assumption than factoring being hard, as far as we know.
It also implies just I-bit messages.
Shorter padding is not known to be secure (at least without a stronger assumption), but is plausibly secure as long as $\ell$ is not too short.

RSA PKCS \#| vl. 5 is a variant of padded RSA where some of the padded bits are random and some have fixed values. With a good choice of parameters, it is probably CPA-secure in the same sense as Padded RSA.

## Another Attack on RSA

Suppose e and m are both small: $m^{e}<N$.

## Another Attack on RSA

Suppose e and m are both small: $m^{e}<N$.
Then the modular arithmetic is irrelevant, since $m^{e} \in \mathbb{Z}_{N}^{*}$ is the same as $m^{e} \in \mathbb{Z}$ (the integers).

## Another Attack on RSA

Suppose e and m are both small: $m^{e}<N$.
Then the modular arithmetic is irrelevant, since $m^{e} \in \mathbb{Z}_{N}^{*}$ is the same as $m^{e} \in \mathbb{Z}$ (the integers).

Here's an attack that takes advantage of this:

## Another Attack on RSA

Suppose e and m are both small: $m^{e}<N$.
Then the modular arithmetic is irrelevant, since $m^{e} \in \mathbb{Z}_{N}^{*}$ is the same as $m^{e} \in \mathbb{Z}$ (the integers).

Here's an attack that takes advantage of this:
I. Guess a value $m_{0}$ for $m$. Compute $m_{0}^{e}$.

## Another Attack on RSA

Suppose e and m are both small: $m^{e}<N$.
Then the modular arithmetic is irrelevant, since $m^{e} \in \mathbb{Z}_{N}^{*}$ is the same as $m^{e} \in \mathbb{Z}$ (the integers).

Here's an attack that takes advantage of this:
I . Guess a value $m_{0}$ for m . Compute $m_{0}^{e}$.
2. If $m_{0}^{e}<m^{e}$, guess a new $m_{1}$ that is larger.

## Another Attack on RSA

Suppose e and m are both small: $m^{e}<N$.
Then the modular arithmetic is irrelevant, since $m^{e} \in \mathbb{Z}_{N}^{*}$ is the same as $m^{e} \in \mathbb{Z}$ (the integers).

Here's an attack that takes advantage of this:
$I$. Guess a value $m_{0}$ for $m$. Compute $m_{0}^{e}$.
2. If $m_{0}^{e}<m^{e}$, guess a new $m_{1}$ that is larger.
3. Otherwise, guess a new smaller $m_{1}$.

## Another Attack on RSA

Suppose e and m are both small: $m^{e}<N$.
Then the modular arithmetic is irrelevant, since $m^{e} \in \mathbb{Z}_{N}^{*}$ is the same as $m^{e} \in \mathbb{Z}$ (the integers).

Here's an attack that takes advantage of this:
I. Guess a value $m_{0}$ for $m$. Compute $m_{0}^{e}$.
2. If $m_{0}^{e}<m^{e}$, guess a new $m_{1}$ that is larger.
3. Otherwise, guess a new smaller $m_{1}$.
4. Repeat steps I-3, narrowing the search space exponentially (for instance, by choosing an $m_{i}$ halfway between the current upper and lower bounds) until $m_{i}^{e}=m^{e}$.

## Another Attack on RSA

Suppose e and m are both small: $m^{e}<N$.
Then the modular arithmetic is irrelevant, since $m^{e} \in \mathbb{Z}_{N}^{*}$ is the same as $m^{e} \in \mathbb{Z}$ (the integers).

Here's an attack that takes advantage of this:
I. Guess a value $m_{0}$ for $m$. Compute $m_{0}^{e}$.
2. If $m_{0}^{e}<m^{e}$, guess a new $m_{1}$ that is larger.
3. Otherwise, guess a new smaller $m_{1}$.
4. Repeat steps I-3, narrowing the search space exponentially (for instance, by choosing an $m_{i}$ halfway between the current upper and lower bounds) until $m_{i}^{e}=m^{e}$.
Why doesn't this work if $m^{e}>N$ ?

## Another Attack on RSA

Suppose e and m are both small: $m^{e}<N$.
Then the modular arithmetic is irrelevant, since $m^{e} \in \mathbb{Z}_{N}^{*}$ is the same as $m^{e} \in \mathbb{Z}$ (the integers).

Here's an attack that takes advantage of this:
I. Guess a value $m_{0}$ for $m$. Compute $m_{0}^{e}$.
2. If $m_{0}^{e}<m^{e}$, guess a new $m_{1}$ that is larger.
3. Otherwise, guess a new smaller $m_{1}$.
4. Repeat steps I-3, narrowing the search space exponentially (for instance, by choosing an $m_{i}$ halfway between the current upper and lower bounds) until $m_{i}^{e}=m^{e}$.
Why doesn't this work if $m^{e}>N$ ?
Because < and > are not well-defined in modular arithmetic!

## Another Attack on RSA

Suppose e and m are both small: $m^{e}<N$.
Then the modular arithmetic is irrelevant, since $m^{e} \in \mathbb{Z}_{N}^{*}$ is the same as $m^{e} \in \mathbb{Z}$ (the integers).

Here's an attack that takes advantage of this:
I. Guess a value $m_{0}$ for $m$. Compute $m_{0}^{e}$.
2. If $m_{0}^{e}<m^{e}$, guess a new $m_{1}$ that is larger.
3. Otherwise, guess a new smaller $m_{1}$.
4. Repeat steps I-3, narrowing the search space exponentially (for instance, by choosing an $m_{i}$ halfway between the current upper and lower bounds) until $m_{i}^{e}=m^{e}$.
Why doesn't this work if $m^{e}>N$ ?
Because < and > are not well-defined in modular arithmetic!
Taking ${ }^{\text {th }}$ roots is easy in $\mathbb{Z}$ and hard in $\mathbb{Z}_{N}^{*}$ !

This class is being recorded

## Picking e

How do we pick e?

- e must be relatively prime to $\varphi(N)$ so that it has an inverse $\bmod \varphi(N)$.
- If e and $m$ are both small ( $m^{e}<N$ ), decrypting is easy.
- But when we pad the high bits of $m$, this is very unlikely, so any e works.
- Might as well pick an e so that it is easy to calculate $m^{e} \bmod N$.
- Using repeated squaring, the \# of terms to multiply together is the \# of I's in the binary representation of e.
- So pick e with weight 2, e.g., e=3.
- Or maybe a larger e of the form $2^{a}+1$ to avoid the possible small e and small mattack.


## Algorithms for Factoring

A factoring algorithm breaks RSA ... so how hard is factoring?
Similar to discrete log:

- When $N=p q$ and $p-I$ is a product of small primes, Pollard's p-I algorithm factors efficiently.
- For general N , the number field sieve runs in time $2^{O\left((\log N)^{1 / 3}(\log \log N)^{2 / 3}\right)}$ (apparently). This is sub-exponential.
- Except:A quantum computer can efficiently factor arbitrary N .

So recommended key lengths for secure variants of RSA are 2048 bits and higher.

## KEM/DEM

Recall that a KEM creates and sends a random key string to someone whose public key you have.

That key then becomes the key for a private-key encryption system, which is here called a data-encapsulation mechanism (DEM)


Alice


Bob

## KEM/DEM

Recall that a KEM creates and sends a random key string to someone whose public key you have.

That key then becomes the key for a private-key encryption system, which is here called a data-encapsulation mechanism (DEM)


Alice



Bob

## KEM/DEM

Recall that a KEM creates and sends a random key string to someone whose public key you have.

That key then becomes the key for a private-key encryption system, which is here called a data-encapsulation mechanism (DEM)


Alice



Bob
message m

## KEM/DEM

Recall that a KEM creates and sends a random key string to someone whose public key you have.

That key then becomes the key for a private-key encryption system, which is here called a data-encapsulation mechanism (DEM)


Alice



Bob
message m

This class is being recorded

## KEM/DEM

Recall that a KEM creates and sends a random key string to someone whose public key you have.

That key then becomes the key for a private-key encryption system, which is here called a data-encapsulation mechanism (DEM)


This class is being recorded

## KEM/DEM

Recall that a KEM creates and sends a random key string to someone whose public key you have.

That key then becomes the key for a private-key encryption system, which is here called a data-encapsulation mechanism (DEM)


This class is being recorded

## KEM/DEM

Recall that a KEM creates and sends a random key string to someone whose public key you have.

That key then becomes the key for a private-key encryption system, which is here called a data-encapsulation mechanism (DEM)


Bob

This class is being recorded

## KEM/DEM

Recall that a KEM creates and sends a random key string to someone whose public key you have.

That key then becomes the key for a private-key encryption system, which is here called a data-encapsulation mechanism (DEM)


This class is being recorded

## RSA-based KEM

Gen: Pick a (public) key derivation function $\mathrm{H}(\mathrm{x})$, then as usual for RSA, i.e., generate two random primes p and q which are $s$ bits long. Let $\mathrm{N}=\mathrm{pq}$. Choose $e, d \in \mathbb{Z}_{\varphi(N)}^{*}$ s.t. $e d=1 \bmod \varphi(N)$. The public key is $(\mathrm{N}, \mathrm{e})$ and the private key is $(\mathrm{N}, \mathrm{d})$.
Encaps: Choose random $x$. The ciphertext is $c=x^{e} \bmod N$ and the key is $\mathrm{H}(\mathrm{x})$.
Decaps: Given c and d , compute $x^{\prime}=c^{d} \bmod N$. Then the key is H(x').

As with Diffie-Hellman-based KEM/DEM, this allows us to combine the strengths of public key and private key cryptography.

This KEM is secure given RSA assumption and an assumption on the key derivation function $\mathrm{H}(\mathrm{x})$.

## How Do We Distribute Public Keys?

Recall that Diffie-Hellman had a man-in-the-middle attack where Eve pretended to be Bob when talking to Alice and vice-versa.

Public key systems have the same sort of problem at the time of distribution of the public keys:


Alice


Bob

## How Do We Distribute Public Keys?

Recall that Diffie-Hellman had a man-in-the-middle attack where Eve pretended to be Bob when talking to Alice and vice-versa.

Public key systems have the same sort of problem at the time of distribution of the public keys:


Alice



Eve

How does Alice know Bob's public key really comes from Bob and not from Eve? Bob needs some way to authenticate it.

## Chosen Ciphertext Attack on RSA

Plain RSA and Padded RSA are vulnerable to another kind of attack with a stronger threat model, a chosen ciphertext attack:

Suppose Alice sends Bob the ciphertext $c=\tilde{m}^{e} \bmod N$.
Eve can easily create $c^{\prime}=2^{e} c=(2 \tilde{m})^{e} \bmod N$.
Why is that a problem? What if Eve sends c' to Bob, who decrypts it and then acts on the message assuming it to be from Alice. By observing Bob, Eve may be able to deduce the plaintext corresponding to $c^{\prime}$, namely $2 \tilde{m}$. This tells her $\tilde{m}$.

Authenticating messages will help us deal with this class of attacks as well, by removing Eve's ability to change the message.

- Message authentication and digital signatures to ensure validity of public keys.
- Hash functions to serve as key derivation functions.
- Chosen-ciphertext attacks (CCA) and CCA-security.

Message authentication, digital signatures, and hash functions all have other applications beyond their usefulness for public key cryptography.

