# CS 7480 Special Topics in PL

Class 2: Foundations Sep 10, 2024



Joshua Gancher

## Today

• Build a knowledge base to help you be successful in this class

Foundations in Provable Cryptography

## Cryptography

Scope for today:

Crypto used in communication protocols: TLS/WireGuard/Kerberos/...

Everything you need to know to understand what we will be reading for Part 1

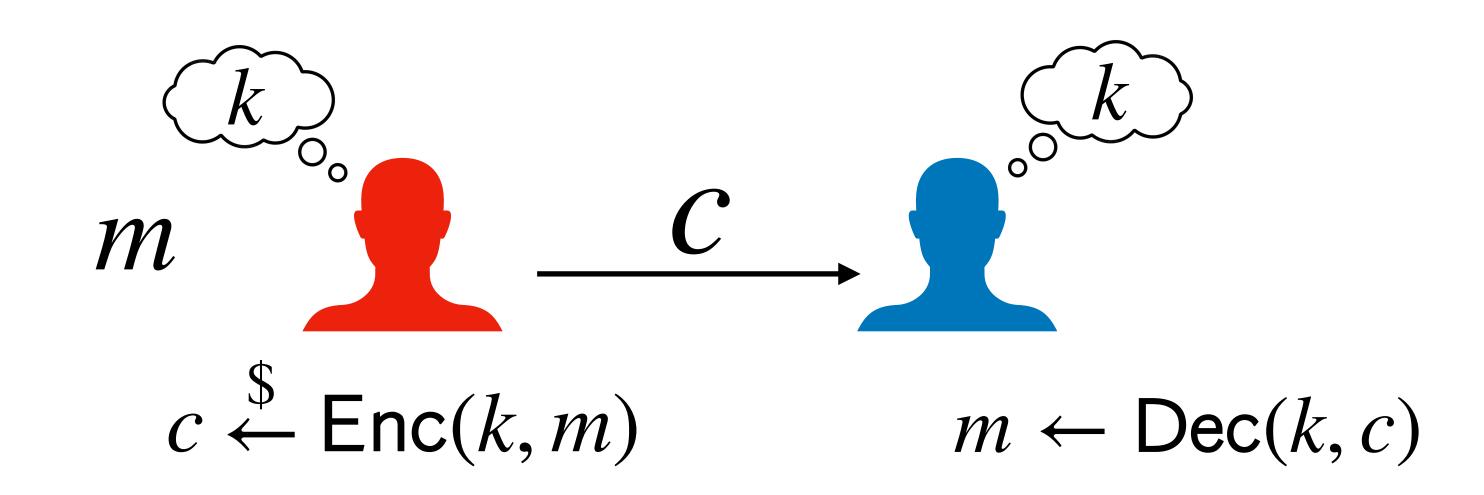
Won't touch on "next-gen crypto": Multi-Party Computation, Blockchains, ZKP, ...

### The Fundamental Problem

- Communication protocols are designed to work over open networks
  - Attacker can read and modify any message, control scheduling

- How to regain any security?
  - Parties must hold secret keys, take advantage of their secrecy

## Key Primitive: Symmetric Encryption



Encryption

Scheme:

Gen : probabilistically outputs a key  $k \in \{0,1\}^*$ 

Enc: given key k and  $m \in \{0,1\}^*$ , probabilistically outputs **ciphertext**  $c \in \{0,1\}^*$ 

Dec: given key k and ciphertext c outputs message or returns an error  $\bot$ 

# Security Properties for Symmetric Encryption

(informally, for Authenticated Encryption)

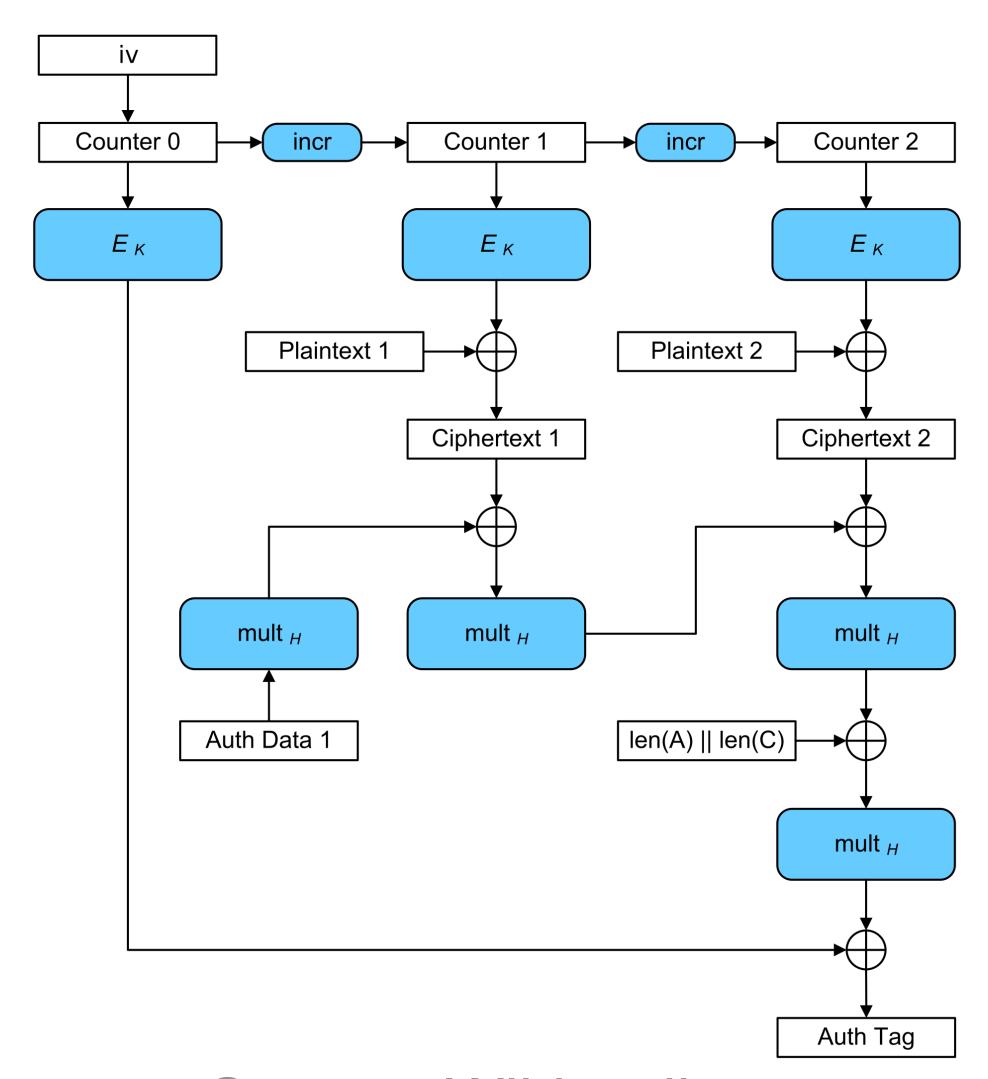
Decryption must be correct:

if k <- Gen and c <- Enc(k, m), then Dec(k, c) = m

If attacker only sees ciphertexts but not otherwise the key:

- Plaintext stays secret, except for its length (Semantic Security)
- Attacker can't produce a valid ciphertext (Ciphertext Integrity)

## Representative Instance: AES-GCM



Source: Wikipedia

Symmetric Crypto:

built on decades of research

into secure ciphers, hashes

security is heuristic (but heavily studied)

Keys are uniformly random bytes

## Main Problem: Key Distribution

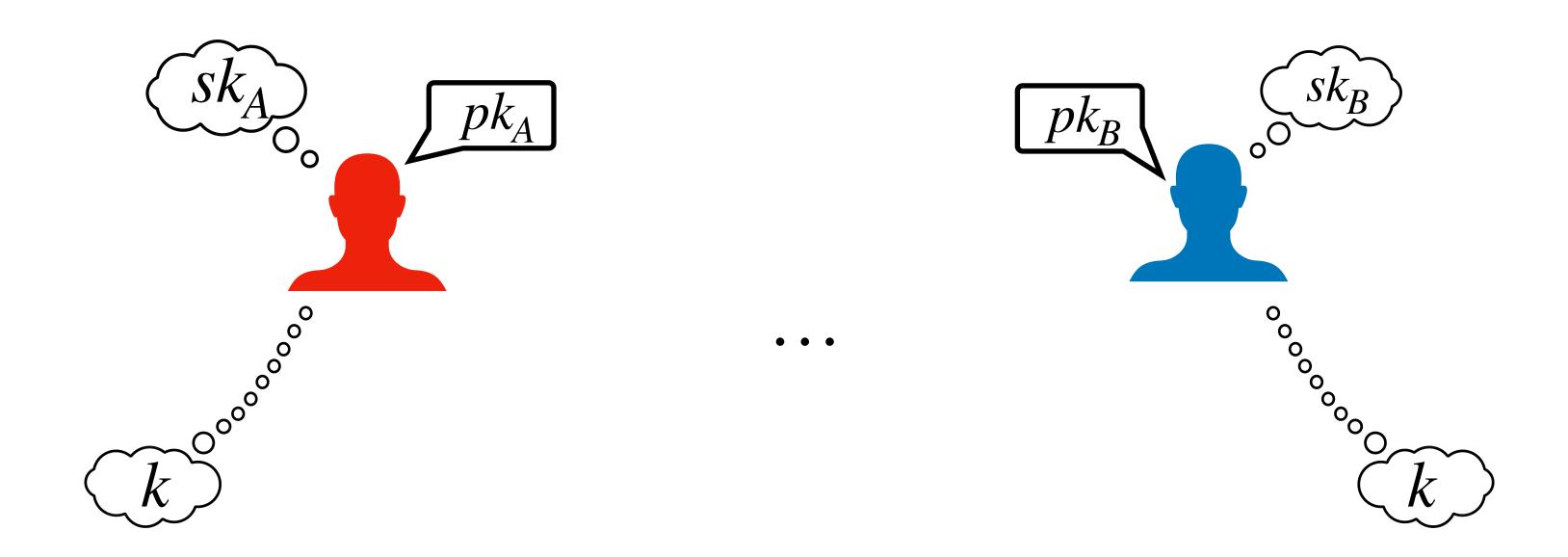
Symmetric encryption only secure if

both parties hold the same secret, randomly generated key

- If key isn't a secret: attacker can decrypt, forge ciphertexts
- If key isn't the same: communication will fail, or parties can be deceived
  - A thinks it's talking to B, but really talking to C
- If key isn't generated correctly, attacker can use this to attack
  - If first bit = 0, now there are half as many possible keys!

# Solutions to Key Generation

- 1. Manually deliver the key to both people
  - Unsuitable for the public internet
- 2. Use more cryptography to deliver the key securely
  - We do this by using public key cryptography (PKC)



## **Crypto Necessary for Communication Protocols**

### Encryption

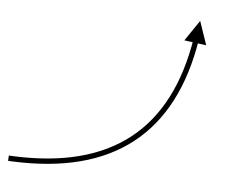
carries out secure comm.

### Key Derivation

converts shared secrets to good encryption keys

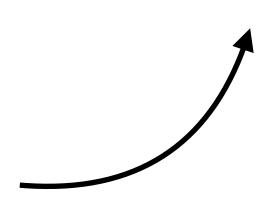
#### Diffie-Hellman

create shared secret using PKC



#### Digital Signatures

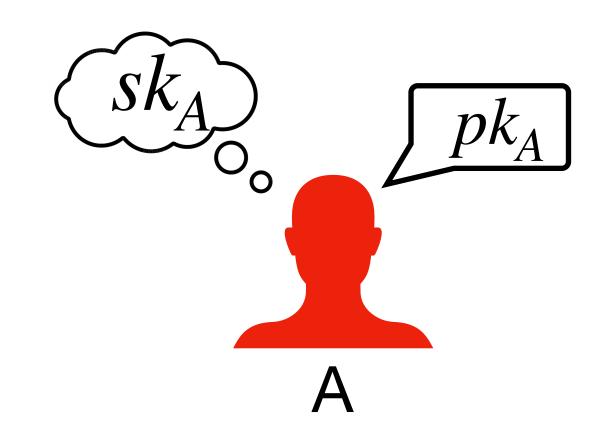
authenticating against public key



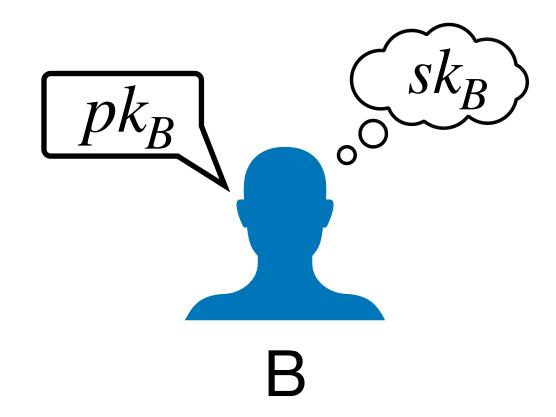
# Diffie-Hellman Exchange











Only B knows B's secret key

Over public channel,

construct shared secret that only A and B know

## Diffie-Hellman Exchange

Gen(): SecretKey

MakePK(sk: SecretKey): PublicKey

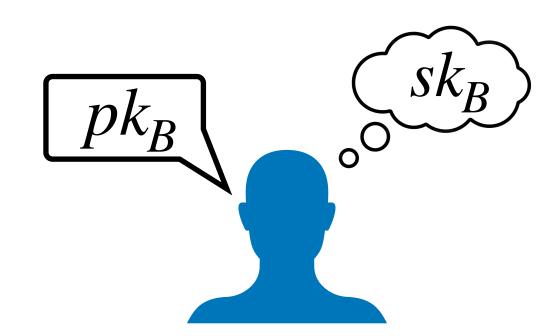
combine(sk : SecretKey, pk : PublicKey) : SharedSecret

my secret key

their public key



 $ss_A := combine(sk_A, pk_B)$ 



 $ss_B := combine(sk_B, pk_A)$ 

# Diffie-Hellman Exchange Properties

correctness: combine( $sk_A$ ,  $pk_B$ ) = combine( $sk_B$ ,  $pk_A$ )

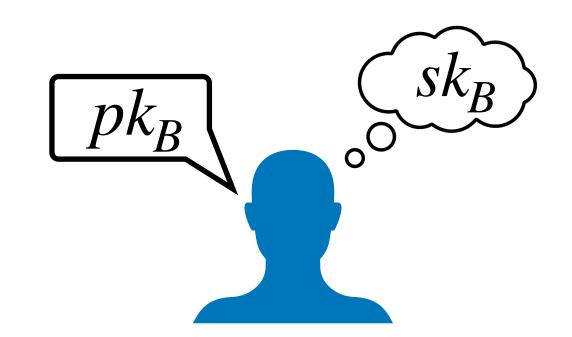
secrecy:

if  $sk_A$ ,  $sk_B$  are kept secret,

SS looks random to attacker



 $ss_A := combine(sk_A, pk_B)$ 



 $ss_B := combine(sk_B, pk_A)$ 

## Diffie-Hellman

under the hood: cyclic groups (often Elliptic Curves)

G cyclic group of order N, generator g

 $Gen() := Uniform(\mathbb{Z}_N)$ 

 $MakePK(x) := g^x$ 

 $combine(x, h) := h^x$ 

combine $(sk_B, pk_A) = (g^b)^a = g^{ba} = (g^a)^b = \text{combine}(sk_A, pk_B)$ 

Diffie-Hellman Assumption:  $g^{ab}$  looks random given  $g^a$ ,  $g^b$ 

## Problem #1: Using the Shared Secret

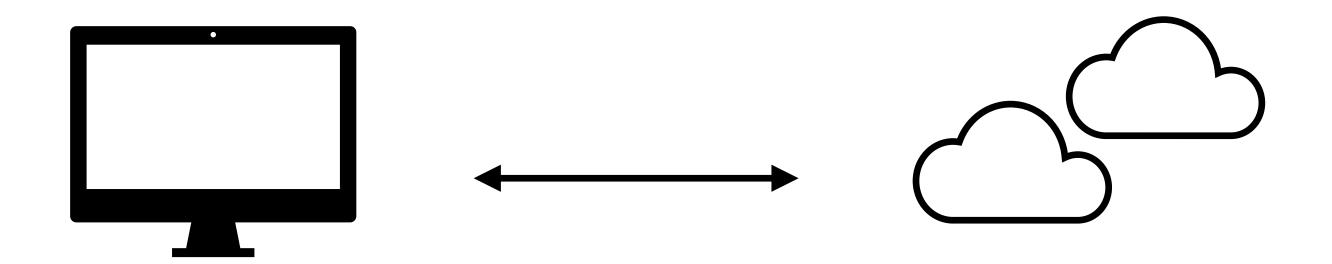
 $ss = combine(sk_B, pk_A)$  has high entropy (unguessable), low uniformity

#### Unsuitable for encryption keys!

Examples: 
$$K' = 010101 \parallel K$$
,  $K$  uniformly random  $K' = K[0] \parallel 0 \parallel K[1] \parallel 0 \parallel ... \parallel K[N]$   $g^{xy}$ , where  $x, y$  uniform

Solution: Key Derivation Function

## Problem #2: Getting the right public key

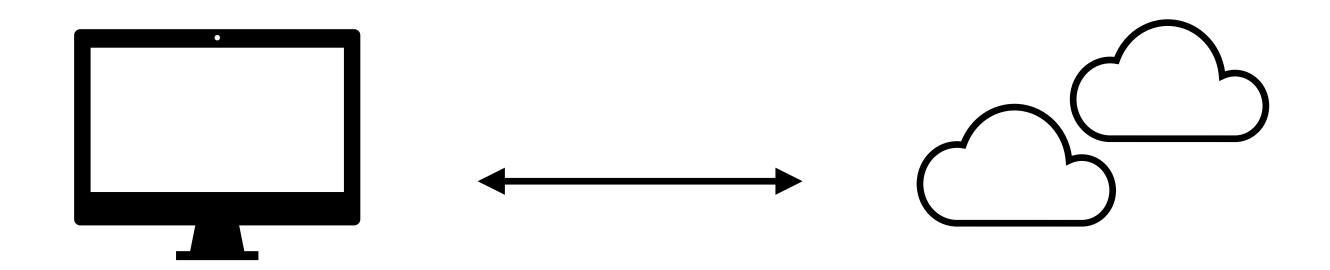


How do I discover Server's public key?

Similar to problem for encryption keys, but big difference:

Distributing public information rather than secret

## Problem #2: Getting the right public key



Option 1: Manually plug in the public key



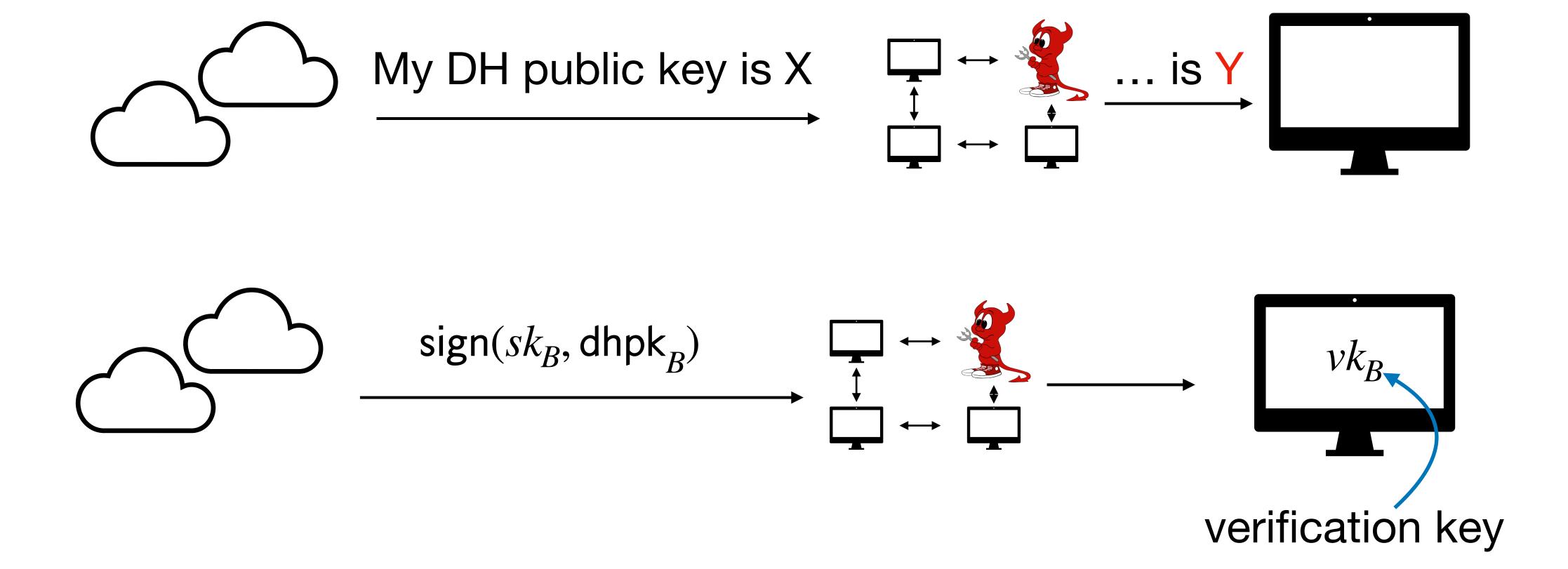
Option 2: Digitally sign the public key with another one

TLS

Allows DH key to change, but signing key stays the same

## Digital Signatures

Proves authenticity of data against public verification key



## Digital Signatures

Gen(): SigningKey

MakeVK(sk: SigningKey): VerifKey

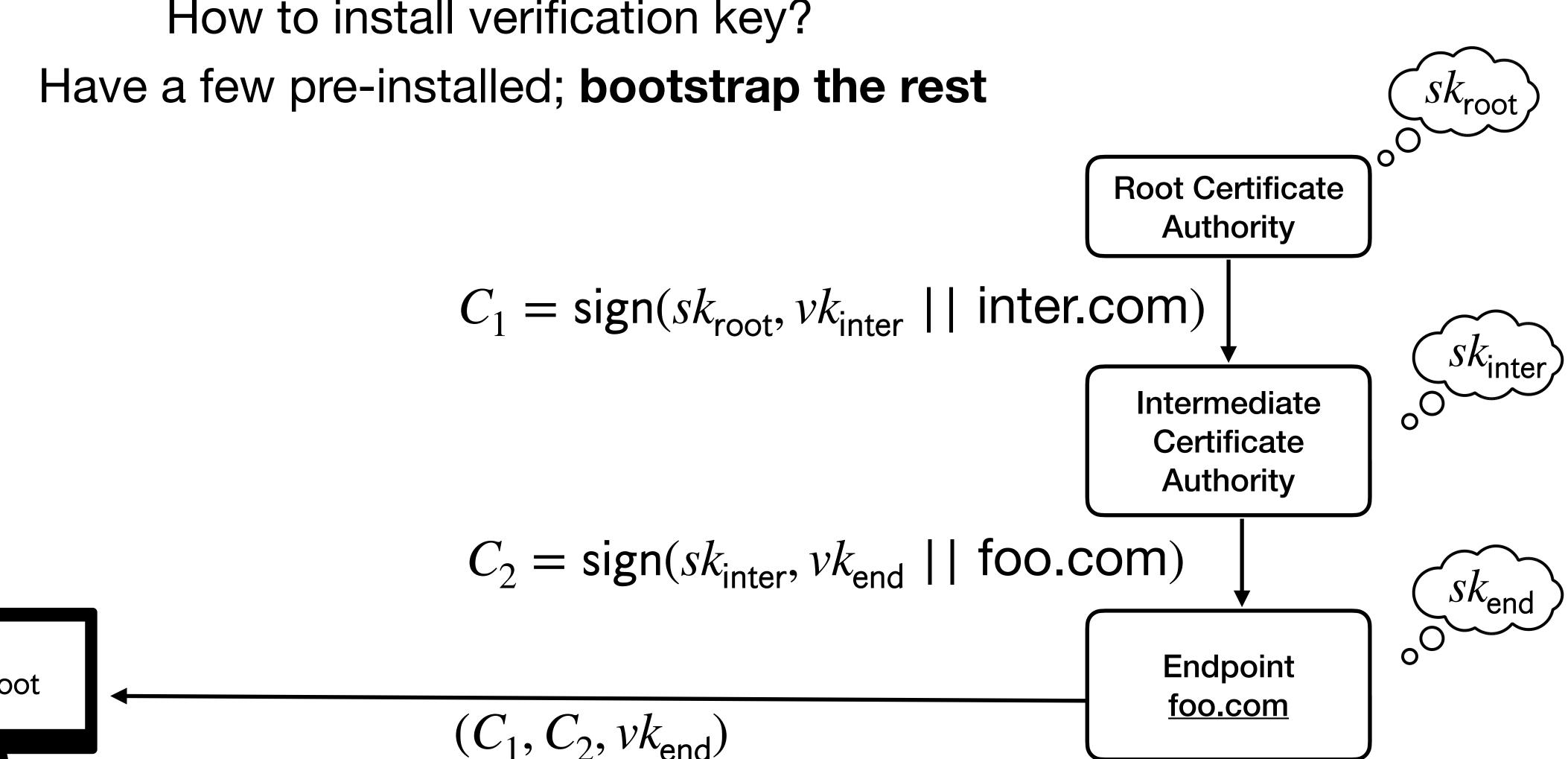
 $Sign(sk : SigningKey, m \in \{0,1\}^*) : Signature$ 

Verify(vk: VerifKey,  $m \in \{0,1\}^*$ , s: Signature): Bool

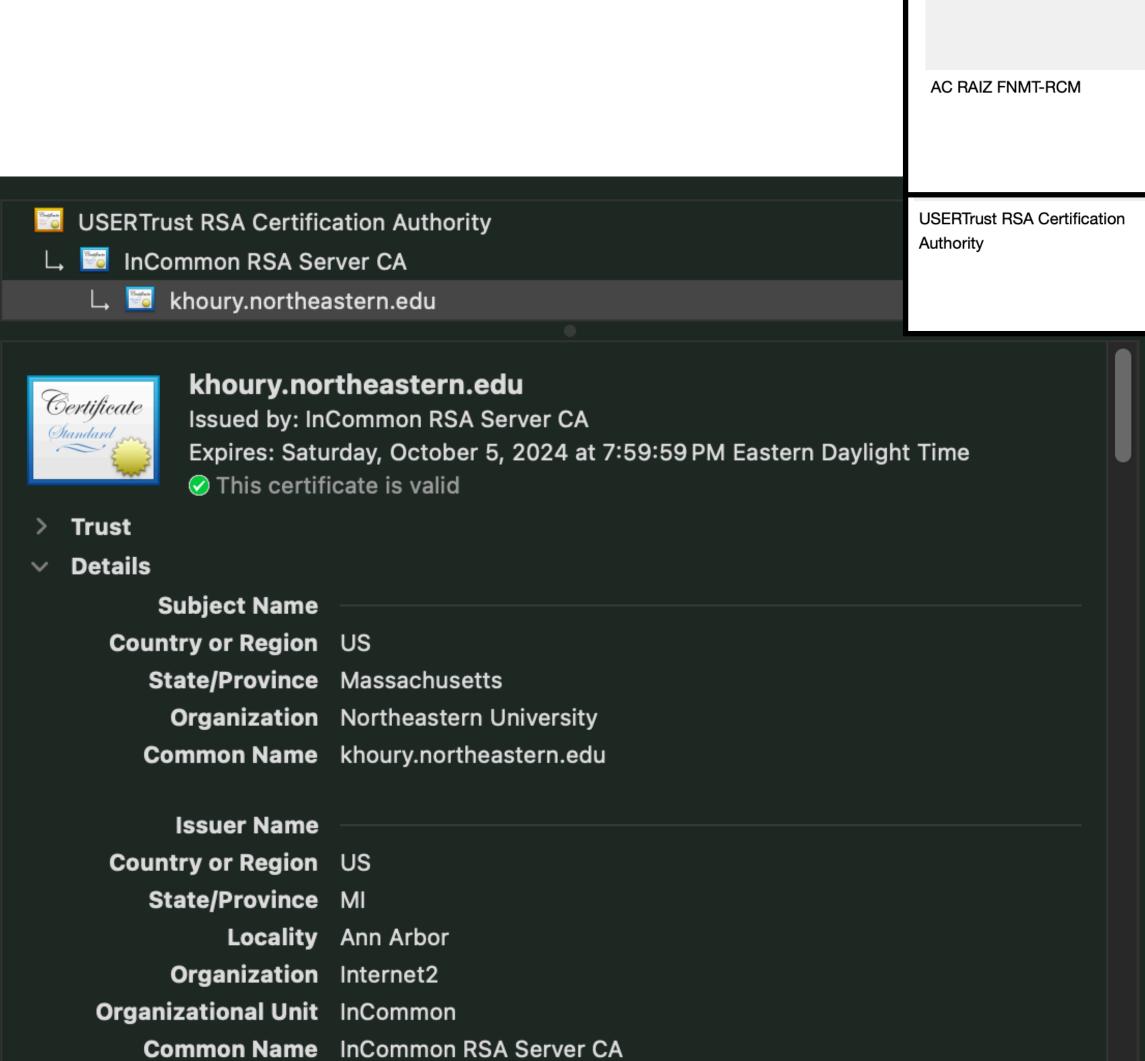
unforgeability: attacker cannot forge signatures using only public verification key, prior signatures

## Getting Correct Verification Keys

How to install verification key?







## **Crypto Necessary for Communication Protocols**

## Encryption

carries out secure comm.

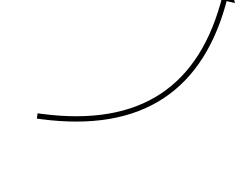
### Key Derivation

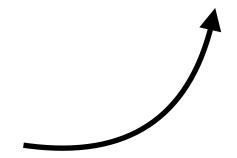
good encryption keys

converts shared secrets to



create shared secret using PKC





#### Digital Signatures

authenticating against public key

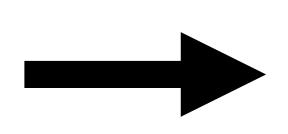
## Specifying and Proving Security

- Up until now: a tour of cryptographic algorithms
- Rest of today: how do we make sure they're used correctly?



Formal specs of security

for encryption, signatures, ...



Formal specs of security

for protocols

## Two Camps for Protocol Security

#### Symbolic Crypto

or, "Dolev-Yao" model

Represent crypto using abstract terms, equations

Closer to existing Formal Methods tools (e.g., Model Checking)

Weaker security guarantees

#### Computational Crypto

Same framework as modern cryptography: complexity theory

Requires more custom tool support

Accurate security guarantees

## Symbolic Crypto

Protocol messages come from BNF grammar

$$t ::= K \in \text{Keys} \mid (t, t) \mid \text{fst}(t) \mid \text{snd}(t) \mid \text{enc}(t, t) \mid \text{dec}(t, t) \mid C \in \text{Const} \mid \dots$$

Example term:

$$fst((enc(K_1, K_2), dec(K_3, 42)))$$

## Symbolic Crypto

Protocol messages come from BNF grammar

$$t ::= K \in \text{Keys} \mid (t, t) \mid \text{fst}(t) \mid \text{snd}(t) \mid \text{enc}(t, t) \mid \text{dec}(t, t) \mid C \in \text{Const} \mid \dots$$

Subject to equations:

$$\operatorname{dec}(k,\operatorname{enc}(k,m)) = m$$
 
$$\operatorname{fst}((x,y)) = x$$
 
$$\operatorname{(reversed argument order from Tamarin)}$$
 
$$\operatorname{snd}((x,y)) = y$$

Attacker allowed to build arbitrary terms using its current knowledge

## Attacker's Knowledge Set

A(t): attacker can deduce t

$$\frac{A(t_1) \ A(t_2)}{A(\mathsf{enc}(t_1, t_2))} \ \frac{A(t_1) \ A(t_2)}{A(\mathsf{dec}(t_1, t_2))} \ \frac{A(t)}{A(\mathsf{fst}(t))} \ \frac{A(t)}{A(\mathsf{snd}(t))} \ \frac{A(C)}{A(C)}$$

If S is a set of terms, let Clo(S) be the closure of S

 $Clo({K_1, K_2}) = all terms built out of <math>K_1, K_2$ 

Example: is  $K_1 \in Clo(\{K_2, enc(K_2, K_1)\})$ ?

## Attacker's Knowledge Set

A(t): attacker can deduce t

$$\frac{A(t_1) \ A(t_2)}{A(\mathsf{enc}(t_1, t_2))} \ \frac{A(t_1) \ A(t_2)}{A(\mathsf{dec}(t_1, t_2))} \ \frac{A(t)}{A(\mathsf{fst}(t))} \ \frac{A(t)}{A(\mathsf{snd}(t))} \ \frac{A(C)}{A(C)}$$

Let 
$$S = \{ \text{enc}(K_3, K_2), \text{enc}(K_2, K_1), \text{enc}(K_2, m), K_3 \}$$
  
Is  $\text{enc}(K_1, m) \in \text{Clo}(S)$ ?

Let 
$$S = \{ \text{enc}(K_3, K_2), \text{enc}(K_2, K_1), \text{enc}(K_2, m), K_3 \}$$
  
Is  $\text{enc}(K_3, m) \in \text{Clo}(S)$ ?

```
output enc(K, N);
recv c;

Alice = if let m = dec(K, c) then
    if m == N + 1 then
    output "ok"
else output "fail"

recv c;

recv c;

if let x = dec(K, c) then
    output enc(K, x + 1)
    else skip
```

Attacker may query parties arbitrarily, subject to its knowledge

Trace event (e.g., input/output) Generates a **trace**  $(A_0,P_0)\overset{e_0}{\to}\dots\overset{e_n}{\to}(A_n,P_n)$  Knowledge of attacker Current state of protocol participant(s)

```
Alice
output enc(K, N);
                                               Attacker Knowledge
recv c;
if let m = dec(K, c) then
                                 enc(K, N)
 if m == N + 1 then
 output "ok"
else output "fail"
      Bob
```

```
recv c;
if let x = dec(K, c) then
  output enc(K, x + 1)
else skip
```

Alice

```
recv c;
                                                   Attacker Knowledge
if let m = dec(K, c) then
                                                      \{\operatorname{enc}(K,N)\}
 if m == N + 1 then
 output "ok"
                           deliver enc(K, N) to Bob
else output "fail"
       Bob
```

```
recv c;
if let x = dec(K, c) then
 output enc(K, x + 1)
else skip
```

#### Alice

```
recv c;
if let m = dec(K, c) then
  if m == N + 1 then
  output "ok"
else output "fail"
```

#### Bob

```
if let x = dec(K, enc(K, N)) then
  output enc(K, x + 1)
else skip
```

# Attacker Knowledge $\{enc(K, N)\}$

Alice

```
recv c;
if let m = dec(K, c) then
  if m == N + 1 then
  output "ok"
else output "fail"
```

enc(K, N + 1)

Bob

output enc(K, N + 1)

Attacker Knowledge

 $\{\operatorname{enc}(K,N)\}$ 

Alice

skip

#### Alice

```
if let m = dec(K, enc(K, N + 1)) then
  if m == N + 1 then
  output "ok"
else output "fail"

Bob
  skip
```

Attacker Knowledge

 $\{enc(K, N), enc(K, N + 1)\}$ 

Alice

output "ok"

Bob

skip

Attacker Knowledge

 $\{enc(K, N), enc(K, N + 1)\}$ 

# **Example Protocol: Properties**

**Secrecy**: if initial attacker knowledge  $A_0 = \{\}$ , can attacker learn K?

$$(A_0, P_0) \rightarrow^* (A_n, P_n) \implies K \notin Clo(A_n)$$

#### **Authentication:**

If  $A_0 = \{\}$ , and Alice output "ok", Bob must have output enc(K, N + 1)

$$(A_0, P_0) \stackrel{T}{\rightarrow} (A_n, \text{done}) \land (\text{Alice : ok}) \in T \implies (\text{Bob : enc}(K, N+1)) \in T$$

### How to Prove Symbolic Properties

#### Tamarin Prover

The Tamarin prover is a security protocol verification tool that supports both falsification (attack finding) and unbounded verification (proving) in the symbolic model. Security protocols are specified as multiset rewriting systems and analyzed with respect to temporal first-order properties.

Tamarin has been successfully used to analyze and support the development of modern security protocols [1,2], including TLS 1.3 [3,4], 5G-AKA [5,6], Noise [7], EMV (Chip-and-pin) [8], and Apple iMessage [9].





### Weaknesses of Symbolic Crypto

- Encryption is assumed to be completely opaque:
  - The only thing attacker can do is manipulate abstract term
  - In the real world, attacker can:
    - View/change individual bits
    - Leverage brute-force attacks
    - View lengths of all messages

•

### Weaknesses of Symbolic Crypto

#### **NEWS**

# 'CRIME' attack abuses SSL/TLS data compression feature to hijack HTTPS sessions

SSL/TLS data compression leaks information that can be used to decrypt HTTPS session cookies, researchers say

C = encrypt(compress(M))

length of C

— (partial) value of M

# Computational Model

- Same model as is assumed in crypto papers
- Cryptography modeled as probabilistic algorithms on bitstrings
- Attacker is arbitrary probabilistic algorithm

- To rule out brute-force attacks:
  - Attacker is polynomial time
  - Security guarantees must hold up to negligible error

# Negligible Functions

Let A be a probabilistic poly-time (PPT) algorithm

Suppose A is trying to forge a ciphertext c so that dec(K, c) succeeds without knowing K or any valid ciphertext for K

#### Trivial attack:

- 1. Guess K by flipping random coins
- 2. Encrypt 0 using K

$$\Pr[A \text{ wins}] = \frac{1}{2^{\lambda}}$$
  $\lambda = \text{security parameter}$  (here, length of key)

# Negligible Functions

Allow adversaries to violate security with probability negligible in  $\lambda$ 

 $\epsilon(\cdot)$  is negligible when for **all** polynomials P, there exists an N,

$$\lambda > N \implies \epsilon(\lambda) < \frac{1}{P(\lambda)}$$

Example: 
$$\frac{1}{2^{\lambda}}$$
 is negligible  $\frac{1}{2^{1}}$ 

$$\frac{\lambda^5}{2^{\lambda}}$$
 is negligible

### Modelling Encryption Computationally

### The Joy of Cryptography

by Mike Rosulek • joyofcryptography.com • 💆

The Joy of Cryptography is a **free** undergraduate-level textbook that introduces students to the fundamentals of provable security.

Get the full PDF (4.1MB)

Latest draft: Jan 3, 2021; 286 pages

# Security Properties for Symmetric Encryption

(informally, for Authenticated Encryption)

Decryption must be correct:

if k <- Gen and c <- Enc(k, m), then Dec(k, c) = m

If attacker only sees ciphertexts but not otherwise the key:

- Plaintext stays secret, except for its length (Semantic Security)
- Attacker can't produce a valid ciphertext (Ciphertext Integrity)

# Security Games

Security specified via indistinguishability of programs (security games)

### EncReal

init:  $K \leftarrow \$ \operatorname{Gen}(\lambda)$ 

Enc(m):

return enc(K, m)

Dec(c):

return dec(K, c)

initialization code

oracles:

input, return bitstrings

everything implicitly parameterized by  $\lambda$ 

### Adversaries for Games

### EncReal

init:  $K \stackrel{\$}{\leftarrow} \text{Gen}(\lambda)$ 

Enc(m):

return enc(K, m)

Dec(c):

return dec(K, c)

probabilistic algorithm A is an adversary for EncGame if:

- it is polytime in  $\lambda$
- it makes queries to Enc, Dec
- eventually, returns a boolean b

 $A\bowtie \mathsf{EncGame}$ : final decision bit output by A, when linked with EncGame

### Indistinguishability of Games

EncReal

 $\mathsf{init}: \ K \overset{\$}{\leftarrow} \mathsf{Gen}(\lambda)$ 

Enc(m):

return enc(K, m)

Dec(c):

return dec(K, c)

EncIdeal

• • •

Enc(m):

• • •

Dec(c):

**Security Game Indistinguishability:** 

No adversary can tell the difference between the two games

### Indistinguishability of Games

EncReal

init:  $K \stackrel{\$}{\leftarrow} \text{Gen}(\lambda)$ 

Enc(m):

return enc(K, m)

Dec(c):

return dec(K, c)

EncIdeal

• • •

Enc(m):

• • •

Dec(c):

**Security Game Indistinguishability:** 

 $\forall$  PPT A,  $|\Pr[A \bowtie EncGame = 1] - \Pr[A \bowtie EncGame' = 1]| <math>\leq \epsilon(\lambda)$  where  $\epsilon$  is negligible

### EncReal

init:  $K \stackrel{\$}{\leftarrow} \text{Gen}(\lambda)$ 

Enc(m):

return enc(K, m)

Dec(c):

return dec(K, c)



### EncIdeal

init:  $K \stackrel{\$}{\leftarrow} \text{Gen}(\lambda)$ log := []

#### Enc(m):

 $c \leftarrow \operatorname{enc}(K, 0^{|m|})$  $\log[c] := m$ 

return c

#### Dec(c):

if log[c] exists: return log[c]else return \( \perp \)

Main idea: attacker only sees junk ciphertexts

Encryption cannot hide lengths, so junk ciphertext must not obscure length

use an **ideal log** to map junk ciphertexts back to real messages

### EncIdeal

init:  $K \stackrel{\$}{\leftarrow} \text{Gen}(\lambda)$   $\log := []$ 

#### Enc(m):

 $c \leftarrow \operatorname{enc}(K, 0^{|m|})$  $\log[c] := m$ 

log[c] := mreturn c

#### Dec(c):

if  $\log[c]$  exists: return  $\log[c]$  else return  $\bot$ 

correctness: decryption is correct by construction

secrecy: junk ciphertexts don't hold message

integrity: attacker cannot create a ciphertext except by calling Enc oracle; forgery impossible

### EncIdeal

 $\begin{array}{c|c} \text{init} : & K \xleftarrow{\$} \text{Gen}(\lambda) \\ & \log := [] \end{array}$ 

#### Enc(m):

 $c \leftarrow \operatorname{enc}(K, 0^{|m|})$  $\log[c] := m$ 

log[c] := m return c

#### Dec(c):

if  $\log[c]$  exists: return  $\log[c]$  else return  $\bot$ 

Gen, enc, dec is a **secure** 

### authenticated encryption scheme

if the two games are

indistinguishable

### EncReal

init :  $K \stackrel{\$}{\leftarrow} \text{Gen}(\lambda)$ 

#### Enc(m):

return enc(K, m)

#### Dec(c):

return dec(K, c)

### EncIdeal

init:  $K \stackrel{\$}{\leftarrow} \text{Gen}(\lambda)$ log := []

#### Enc(m):

 $c \leftarrow \operatorname{enc}(K, 0^{|m|})$   $\log[c] := m$ 

return c

#### Dec(c):

if log[c] exists: return log[c]else return \( \perp \)

Often phrased as two separate equivalences:

- INT-CTXT (Ciphertext Integrity)
  - Ciphertexts are unforgeable
- IND-CPA (Semantic Security)
  - Ciphertexts keep messages secret

### EncReal

 $\mathsf{init}: K \overset{\$}{\leftarrow} \mathsf{Gen}(\lambda)$ 

Enc(m):

return enc(K, m)

Dec(c):

return dec(K, c)

### EncIdeal

init:  $K \leftarrow \mathbb{G}en(\lambda)$  $\log := []$ 

Enc(m):

 $c \leftarrow \operatorname{enc}(K,0^{|m|})$ 

log[c] := m

return c

Dec(c):

if log[c] exists:

return log[c]

else return \( \perp \)

PReal

```
Alice = output enc(K, N);
```

```
Bob = \begin{cases} recv c; \\ if let x = dec(K, c) then \\ output enc(K, x + 1) \end{cases}
else skip
```

Goal: attacker can't learn anything about N

sim has no access to

simulator

secrets at all!

Goal: 
$$\forall A, \exists S, A \bowtie P_{Real} = S$$

adversary

adv interacting

with secrets

#### Simulation:

"any information that can be computed using protocol can be computed without the protocol"

PReal

```
Alice = output enc(K, N);
```

```
Bob = \begin{cases} recv c; \\ if let x = dec(K, c) then \\ output enc(K, x + 1) \end{cases}
else skip
```

Idea: rewrite protocol to make use of security game for enc, dec

```
Alice = output G.Enc(N);
```

```
Bob = \begin{cases} recv c; \\ if let x = G.Dec(c) then \\ output G.Enc(x + 1) \end{cases}
else skip
```

Idea: rewrite protocol to make use of security game for enc, dec

$$P_{\text{Real}} \approx P_{\text{Hybrid}} \bowtie \text{EncReal}$$

by unfolding definitions

EncReal ≈ EncIdeal

by assumption

 $P_{\text{Real}} \approx P_{\text{Hybrid}} \bowtie \text{EncIdeal}$ 

by congruence

# Hybrid Protocol

# PHybrid

```
Alice = output G.Enc(N);
```

```
Bob = \begin{cases} recv c; \\ if let x = G.Dec(c) then \\ output G.Enc(x + 1) \end{cases}
else skip
```

### Idealized Protocol

### no dataflow from N to output!

$$P_{\mathsf{Ideal}} := P_{\mathsf{Hybrid}} \bowtie \mathsf{EncIdeal}$$

```
log := [];
```

```
let c_out = enc(K, 0000000);
Alice = log[c_out] := N;
    output c_out;
```

### Idealized Protocol

### no dataflow from N to output!

# P'Ideal

log := [];

```
let c_out = enc(K, 0000000);
Alice = log[c_out] := 00000000;
output c_out;
```

### Idealized Protocol

no dataflow from N to output!

$$P_{\mathsf{Ideal}} \approx P_{\mathsf{Ideal}}$$

**Computational Noninterference:** 

Changing value of secret nonce N does not change external behavior

# Proving Secrecy

Goal: 
$$\forall A, \exists S, A \bowtie P_{Real} = S$$

$$A\bowtie P_{\mathsf{Real}} pprox A\bowtie (P_{\mathsf{Hybrid}}\bowtie \mathsf{EncReal})$$
  $pprox A\bowtie (P_{\mathsf{Hybrid}}\bowtie \mathsf{EncIdeal})$   $pprox A\bowtie P_{\mathsf{Ideal}}$   $pprox A\bowtie P'_{\mathsf{Ideal}}=:\mathsf{S}$ 

# Computational Provers

CryptoVerif:

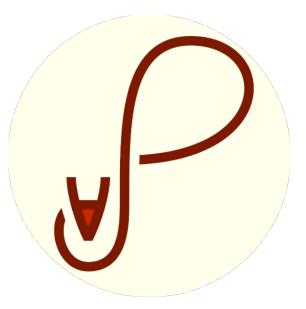
automates security game rewrites



used for TLS, Kerberos, SSH, Signal, WireGuard, ...

#### Squirrel:

hides complex rewrites behind symbolic-looking logic



Owl:

hides complex rewrites behind information flow type system

### Computational Provers

### EasyCrypt:

more expressive, based on Probabilistic Relational Hoare Logic

# Today

• Symbolic Crypto: modelling crypto with abstract terms

Computational Crypto: modelling crypto using indistinguishabilities

### Next Time

### A Comprehensive Symbolic Analysis of TLS 1.3

Cas Cremers University of Oxford, UK Marko Horvat MPI-SWS, Germany Jonathan Hoyland Royal Holloway, University of London, UK

Sam Scott Royal Holloway, University of London, UK Thyla van der Merwe Royal Holloway, University of London, UK

### After that..

# Computationally Sound Mechanized Proofs for Basic and Public-key Kerberos

B. Blanchet \*
CNRS & École Normale
Supérieure & INRIA
blanchet@di.ens.fr

A.D. Jaggard

DIMACS

Rutgers University
adj@dimacs.rutgers.edu

J.-K. Tsay
Department of Mathematics
University of Pennsylvania
jetsay@math.upenn.edu

A. Scedrov<sup>†</sup>
Department of Mathematics
University of Pennsylvania
scedrov@math.upenn.edu