

15-853: Algorithms in the Real World

Cryptography 3 and 4

15-853

Page 1

Cryptography Outline

Introduction: terminology, cryptanalysis, security

Primitives: one-way functions, trapdoors, ...

Protocols: digital signatures, key exchange, ..

Number Theory: groups, fields, ...

Private-Key Algorithms: Rijndael, DES

➔ **Public-Key Algorithms:**

- Diffie-Hellman Key Exchange
- RSA, El-Gamal, Blum-Goldwasser
- Quantum Cryptography

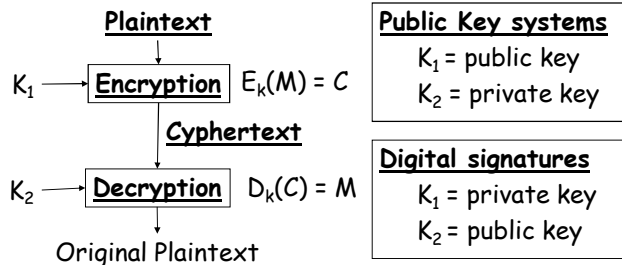
Case Studies: Kerberos, Digital Cash

15-853

Page 2

Public Key Cryptosystems

Introduced by Diffie and Hellman in 1976.



Typically used as part of a more complicated protocol.

15-853

Page 3

One-way trapdoor functions

Both Public-Key and Digital signatures make use of one-way trapdoor functions.

Public Key:

- Encode: $c = f(m)$
- Decode: $m = f^{-1}(c)$ using trapdoor

Digital Signatures:

- Sign: $c = f^{-1}(m)$ using trapdoor
- Verify: $m = f(c)$

15-853

Page 4

Example of SSL (3.0)

SSL (Secure Socket Layer) is the standard for the web (https).

Protocol (somewhat simplified): Bob → amazon.com

B→A: client hello : protocol version, acceptable ciphers	}	hand-shake
A→B: server hello : cipher, session ID, amazon.com _{verisign}		
B→A: key exchange : {masterkey} _{amazon's public key}	}	data
A→B: server finish : ([amazon,prev-messages,masterkey]) _{key1}		
B→A: client finish : ([bob,prev-messages,masterkey]) _{key2}		
A→B: server message : (message1,[message1]) _{key1}		
B→A: client message : (message2,[message2]) _{key2}		

|h|_{issuer} = Certificate
 = Issuer, <h,h's public key, time stamp>_{issuer's private key}
 <...>_{private key} = Digital signature {...}_{public key} = Public-key encryption
 [...] = Secure Hash (...)_{key} = Private-key encryption
 key1 and key2 are derived from masterkey and session ID

15-853

Page 5

Public Key History

Some algorithms

- Diffie-Hellman, 1976, key-exchange based on discrete logs
- Merkle-Hellman, 1978, based on "knapsack problem"
- McEliece, 1978, based on algebraic coding theory
- RSA, 1978, based on factoring
- Rabin, 1979, security can be reduced to factoring
- ElGamal, 1985, based on discrete logs
- Blum-Goldwasser, 1985, based on quadratic residues
- Elliptic curves, 1985, discrete logs over Elliptic curves
- Chor-Rivest, 1988, based on knapsack problem
- NTRU, 1996, based on Lattices
- XTR, 2000, based on discrete logs of a particular field

15-853

Page 6

Diffie-Hellman Key Exchange

A group $(G, *)$ and a primitive element (generator) g is made public.

- Alice picks a , and sends g^a to Bob
- Bob picks b and sends g^b to Alice
- The shared key is g^{ab}

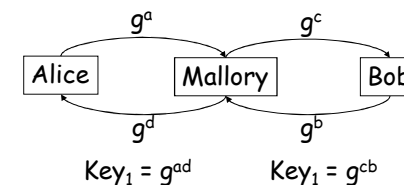
Note this is easy for Alice or Bob to compute, but assuming discrete logs are hard is hard for anyone else to compute.

Can someone see a problem with this protocol?

15-853

Page 7

Person-in-the-middle attack



Mallory gets to listen to everything.

15-853

Page 8

Merkle-Hellman

Gets "security" from the **Subset Sum** (also called **knapsack**) which is NP-hard to solve in general.

Subset Sum (Knapsack): Given a sequence $W = \{w_0, w_1, \dots, w_{n-1}\}$, $w_i \in \mathbb{Z}$ of weights and a sum S , calculate a boolean vector B , such that:

$$\sum_{i=0}^{i < n} B_i w_i = S$$

Even deciding if there is a solution is NP-hard.

Merkle-Hellman

W is **superincreasing** if: $w_i \geq \sum_{j=0}^{i-1} w_j$

It is easy to solve the subset-sum problem for superincreasing W in $O(n)$ time.

Main idea of Merkle-Hellman:

- Hide the easy case by multiplying each w_i by a constant a modulo a prime p
 $w'_i = a * w_i \text{ mod } p$
- Knowing a and p allows you to retrieve the superincreasing sequence

Merkle-Hellman

What we need

- w_1, \dots, w_n
superincreasing integers
- $p > \sum_{i=1}^n w_i$ and prime
- $a, 1 \leq a \leq n$
- $w'_i = a w_i \text{ mod } p$

Public Key: w'_i

Private Key: $w_i, p, a,$

Encode:

$$y = E(m) = \sum_{i=1}^n m_i w'_i$$

Decode:

$$\begin{aligned} z &= a^{-1} y \text{ mod } p \\ &= a^{-1} \sum_{i=1}^n m_i w'_i \text{ mod } p \\ &= a^{-1} \sum_{i=1}^n m_i a w_i \text{ mod } p \\ &= \sum_{i=1}^n m_i w_i \end{aligned}$$

Solve subset sum prob:

(w_1, \dots, w_n, z)

obtaining m_1, \dots, m_n

Merkle Hellman: Problem

Was broken by Shamir in 1984.

Shamir showed how to use integer programming to solve the particular class of Subset Sum problems in polynomial time.

Lesson: don't leave your trapdoor loose.

RSA

Invented by Rivest, Shamir and Adleman in 1978

Based on **difficulty of factoring**.

Used to **hide** the size of a group Z_n^* since:

$$|Z_n^*| = \phi(n) = n \prod_{p|n} (1 - 1/p)$$

Factoring has not been reduced to RSA

- an algorithm that generates m from c does not give an efficient algorithm for factoring

On the other hand, factoring has been reduced to finding the private-key.

- there is an efficient algorithm for factoring given one that can find the private key.

15-853

Page 13

RSA Public-key Cryptosystem

What we need:

- p and q , primes of approximately the same size
- $n = pq$
 $\phi(n) = (p-1)(q-1)$
- $e \in Z_{\phi(n)}^*$
- $d = e^{-1} \bmod \phi(n)$

Public Key: (e, n)

Private Key: d

Encode:

$m \in Z_n$

$$E(m) = m^e \bmod n$$

Decode:

$$D(c) = c^d \bmod n$$

15-853

Page 14

RSA continued

Why it works:

$$D(c) = c^d \bmod n$$

$$= m^{ed} \bmod n$$

$$= m^{1+k(p-1)(q-1)} \bmod n$$

$$= m^{1+k\phi(n)} \bmod n$$

$$= m(m^{\phi(n)})^k \bmod n$$

$$= m$$

Why is this argument not quite sound?

What if $m \notin Z_n^*$ then $m^{\phi(n)} \neq 1 \bmod n$

Answer 1: Not hard to show that it still works.

Answer 2: jackpot - you've factored n

15-853

Page 15

RSA computations

To **generate the keys**, we need to

- **Find two primes p and q .** Generate candidates and use primality testing to filter them.
- **Find $e^{-1} \bmod (p-1)(q-1)$.** Use Euclid's algorithm. Takes time $\log^2(n)$

To **encode and decode**

- **Take m^e or c^d .** Use the power method. Takes time $\log(e) \log^2(n)$ and $\log(d) \log^2(n)$.

In practice e is selected to be small so that encoding is fast.

15-853

Page 16

Security of RSA

Warning:

- Do not use this or any other algorithm naively!

Possible security holes:

- Need to use "safe" primes p and q. In particular p-1 and q-1 should have large prime factors.
- p and q should not have the same number of digits. Can use a middle attack starting at sqrt(n).
- e cannot be too small
- Don't use same n for different e's.
- You should always "pad"

Algorithm to factor given d and e

If an attacker has an algorithm that generates d from e, then he/she can factor n in PPT. Variant of the Rabin-Miller primality test.

Function TryFactor(e,d,n)

1. write $ed - 1$ as $2^s r$, r odd
2. choose w at random $< n$
3. $v = w^r \pmod n$
4. if $v = 1$ then return(**fail**)
5. while $v \neq 1 \pmod n$
6. $v_0 = v$
7. $v = v^2 \pmod n$
8. if $v_0 = n - 1$ then return(**fail**)
9. return(**pass**, gcd($v_0 + 1$, n))

LasVegas algorithm
Probability of pass is $> .5$.
Will return p or q if it passes.
Try until you pass.

RSA Performance

Performance: (600Mhz PIII) (from: [ssh toolkit](#)):

Algorithm	Bits/key		Mbits/sec
RSA Keygen	1024	.35sec/key	
	2048	2.83sec/key	
RSA Encrypt	1024	1786/sec	3.5
	2048	672/sec	1.2
RSA Decrypt	1024	74/sec	.074
	2048	12/sec	.024
ElGamal Enc.	1024	31/sec	.031
ElGamal Dec.	1024	61/sec	.061
DES-cbc	56		95
twofish-cbc	128		140
Rijndael	128		180

RSA in the "Real World"

Part of many standards: PKCS, ITU X.509, ANSI X9.31, IEEE P1363

Used by: SSL, PEM, PGP, Entrust, ...

The standards specify many details on the implementation, e.g.

- e should be selected to be small, but not too small
- "multi prime" versions make use of $n = pqr...$ this makes it cheaper to decode especially in parallel (uses Chinese remainder theorem).

Factoring in the Real World

Quadratic Sieve (QS): (Here, n is the number to factor, not # of bits to encode it)

$$T(n) = e^{(1+o(1))(\ln n)^{1/2} (\ln(\ln n))^{1/2}}$$

- Used in 1994 to factor a 129 digit (428-bit) number. 1600 Machines, 8 months.

Number field Sieve (NFS):

$$T(n) = e^{(1.923+o(1))(\ln n)^{1/3} (\ln(\ln n))^{2/3}}$$

- Used in 1999 to factor 155 digit (512-bit) number. 35 CPU years. At least 4x faster than QS
- Used in 2003-2005 to factor 200 digits (663 bits) 75 CPU years (\$20K prize)

15-853

Page 21

ElGamal

Based on the difficulty of the discrete log problem.

Invented in 1985

Digital signature and Key-exchange variants

- Digital signature is AES standard
- Public Key used by TRW (avoided RSA patent)

Works over various groups

- Z_p ,
- Multiplicative group $GF(p^n)$,
- Elliptic Curves

15-853

Page 22

ElGamal Public-key Cryptosystem

$(G, *)$ is a group

- α a generator for G
- $a \in Z_{|G|}$
- $\beta = \alpha^a$

G is selected so that it is hard to solve the discrete log problem.

Public Key: (α, β) and some description of G

Private Key: a

Encode:

Pick random $k \in Z_{|G|}$

$$E(m) = (y_1, y_2) \\ = (\alpha^k, m * \beta^k)$$

Decode:

$$D(y) = y_2 * (y_1^a)^{-1} \\ = (m * \beta^k) * (\alpha^{ka})^{-1} \\ = m * \beta^{k * (\beta^k)^{-1}} \\ = m$$

You need to know a to easily decode $y!$

15-853

Page 23

ElGamal: Example

$G = Z_{11}^*$

- $\alpha = 2$
- $a = 8$
- $\beta = 2^8 \pmod{11} = 3$

Encode: 7

Pick random $k = 4$

$$E(m) = (2^4, 7 * 3^4) \\ = (5, 6)$$

Decode: (5, 6)

$$D(y) = 6 * (5^8)^{-1} \\ = 6 * 4^{-1} \\ = 6 * 3 \pmod{11} \\ = 7$$

Public Key: $(2, 3), Z_{11}^*$

Private Key: $a = 8$

15-853

Page 24

Probabilistic Encryption

For RSA one message goes to one cipher word. This means we might gain information by running $E_{\text{public}}(M)$.

Probabilistic encryption maps every M to many C randomly. Cryptanalysts can't tell whether $C = E_{\text{public}}(M)$.

ElGamal is an example (based on the random k), but it doubles the size of message.

15-853

Page 25

BBS "secure" random bits

BBS (Blum, Blum and Shub, 1984)

- Based on difficulty of factoring, or finding square roots modulo $n = pq$.

Fixed

- p and q are primes such that $p = q = 3 \pmod{4}$
- $n = pq$ (is called a Blum integer)

For a particular bit seq.

- **Seed:** random x relatively prime to n .
- **Initial state:** $x_0 = x^2$
- **i^{th} state:** $x_i = (x_{i-1})^2$
- **i^{th} bit:** lsb of x_i

Note that: $x_0 = x_i^{-2^i \pmod{\phi(n)}} \pmod{n}$

Therefore knowing p and q allows us to find x_0 from x_i

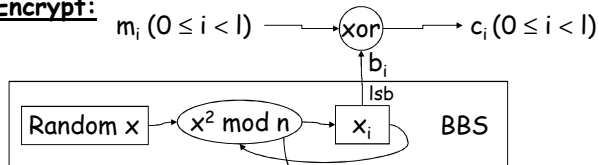
15-853

Page 26

Blum-Goldwasser: A stream cipher

Public key: $n (= pq)$ **Private key:** p or q

Encrypt: $m_i (0 \leq i < l) \xrightarrow{\text{xor}} c_i (0 \leq i < l)$



Last $\log(n)$ bits are x_i : $c_i (l \leq i < l + \log n) = x_i$

Decrypt:

Using p and q , find $x_0 = x_i^{-2^i \pmod{(p-1)(q-1)}} \pmod{n}$

Use this to regenerate the b_i and hence m_i

15-853

Page 27

Quantum Cryptography

In quantum mechanics, there is no way to take a measurement without potentially changing the state. E.g.

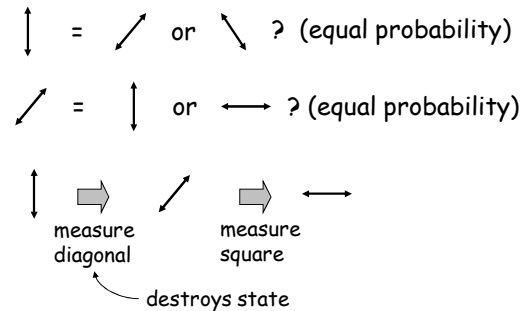
- Measuring position, spreads out the momentum
- Measuring spin horizontally, "spreads out" the spin probability vertically

Related to Heisenberg's uncertainty principal

15-853

Page 28

Using photon polarization



15-853

Page 29

Quantum Key Exchange

1. Alice sends bob photon stream randomly polarized in one of 4 polarizations: $\updownarrow \leftarrow \rightarrow \nearrow \searrow$
2. Bob measures photons in random orientations
 e.g.: $\times + + \times \times \times + \times$ (orientations used)
 $\backslash \mid - \backslash / / - \backslash$ (measured polarizations)
 and tells Alice in the open what orientations he used, but not what he measured.
3. Alice tells Bob in the open which are correct
4. Bob and Alice keep the correct values
 Susceptible to a **man-in-the-middle** attack

15-853

Page 30

In the "real world"

DARPA Quantum Network: 10 node network (BBN, Harvard, B.U.)

Los Alamos/NIST demo: ~150 km of fiber (2007)

Commercial products: id Quantique, MagiQ, SmartQuantim



15-853

Page 31

Cryptography Outline

Introduction: terminology, cryptanalysis, security

Primitives: one-way functions, trapdoors, ...

Protocols: digital signatures, key exchange, ..

Number Theory: groups, fields, ...

Private-Key Algorithms: Rijndael, DES

Public-Key Algorithms: Knapsack, RSA, El-Gamal, ...

→ **Case Studies:**

- Kerberos
- Digital Cash

15-853

Page 32

Kerberos

A key-serving system based on Private-Keys (DES).

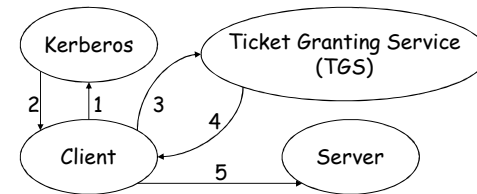
Assumptions

- Built on top of TCP/IP networks
- Many "**clients**" (typically users, but perhaps software)
- Many "**servers**" (e.g. file servers, compute servers, print servers, ...)
- User machines and servers are potentially insecure without compromising the whole system
- A **kerberos server** must be secure.

15-853

Page 33

Kerberos



1. Request *ticket-granting-ticket* (TGT)
2. <TGT>
3. Request *server-ticket* (ST)
4. <ST>
5. Request service

15-853

Page 34

Kerberos V Message Formats

C = client S = server K = key

T = timestamp V = time range

TGS = Ticket Granting Service A = Net Address

Ticket Granting Ticket: $T_{C,TGS} = TGS, \{C, A, V, K_{C,TGS}\} K_{TGS}$

Server Ticket: $T_{C,S} = S, \{C, A, V, K_{C,S}\} K_S$

Authenticator: $A_{C,S} = \{C, T, [K]\} K_{C,S}$

1. Client to Kerberos: $\{C, TGS\} K_C$
 2. Kerberos to Client: $\{K_{C,TGS}\} K_C, T_{C,TGS}$
 3. Client to TGS: $A_{C,TGS}, T_{C,TGS}$
 4. TGS to Client: $\{K_{C,S}\} K_{C,TGS}, T_{C,S}$
 5. Client to Server: $A_{C,S}, T_{C,S}$
- ← Possibly repeat

15-853

Page 35

Kerberos Notes

All machines have to have synchronized clocks

- Must not be able to reuse authenticators

Servers should store all previous and valid tickets

- Help prevent replays

Client keys are typically a one-way hash of the password. Clients do not keep these keys.

Kerberos 5 uses CBC mode for encryption Kerberos 4 was insecure because it used a nonstandard mode.

15-853

Page 36

Electronic Payments

Privacy

- Identified
- Anonymous

Involvement

- Offline (just buyer and seller)
more practical for "micropayments"
- Online
 - Notational fund transfer (e.g. Visa, CyberCash)
 - Trusted 3rd party (e.g. FirstVirtual)

Today: "Digital Cash" (anonymous and possibly offline)

15-853

Page 37

Some more protocols

1. Secret splitting (and sharing)
2. Bit commitment
3. Blind signatures

15-853

Page 38

Secret Splitting

Take a secret (e.g. a bit-string **B**) and split it among multiple parties such that all parties have to cooperate to regenerate any part of the secret.

An implementation:

- Trent picks a random bit-string **R** of same length as **B**
- Sends Alice **R**
- Sends Bob **R xor B**

Generalizes to k parties by picking $k-1$ random values.

15-853

Page 39

Secret Sharing

m out of n ($m < n$) parties can recreate the secret.

Also called an **(m, n)-threshold scheme**

An implementation (Shamir):

- Write secret as coefficients of a polynomial $GF(p^l)[x]$ of order $m-1$ ($n \leq p^l$).
$$p(x) = c_{m-1}x^{m-1} + \dots + c_1 x + c_0$$
- Evaluate $p(x)$ at n distinct points in $GF(p^l)$
- Give each party one of the results
- Any m results can be used to reconstruct the polynomial.

15-853

Page 40

Bit Commitment

Alice commits a bit to Bob without revealing the bit (until Bob asks her to prove it later)

An implementation:

- **Commit**
 - Alice picks random r , and uses a one-way hash function to generate $y = f(r, b)$
 - $f(r, b)$ must be "unbiased" on b (y by itself tells you nothing about b).
 - Alice **sends** Bob y .
- **Open** (expose bit and prove it was committed)
 - Alice **sends** Bob b and r .

Example: $y = \text{Rijndael}_r(000\dots b)$

Blind Signatures

Sign a message m without knowing anything about m
Sounds dangerous, but can be used to give "value" to an anonymous message

- Each signature has meaning:
\$5 signature, \$20 signature, ...

Blind Signatures

An implementation: based on RSA

Trent blindly signs a message m from Alice

- Trent has public key (e, n) and private key d
- Alice selects random $r < n$ and generates
 $m' = m r^e \text{ mod } n$
and sends it to Trent.
This is called **blinding** m

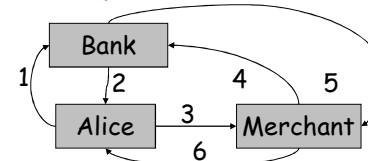
- Trent signs it: $s(m') = (m r^e)^d \text{ mod } n$

- Alice calculates:

$$s(m) = s(m') r^{-1} = m^d r^{ed-1} = m^d \text{ mod } n$$

Patented by Chaum in 1990.

An anonymous online scheme



1. Blinded Unique Random large ID (no collisions).
 $\text{Sig}_{\text{alice}}(\text{request for } \$100)$.
2. $\text{Sig}_{\text{bank_}\$100}(\text{blinded}(\text{ID}))$: signed by bank, Alice debited
3. $\text{Sig}_{\text{bank_}\$100}(\text{ID})$
4. $\text{Sig}_{\text{bank_}\$100}(\text{ID})$
5. OK from bank (bank stores ID)
6. OK from merchant

Minting: 1. and 2.
Spending: 3-6.
Left out encryption for simplicity.

eCash

Uses the protocol

Bought assets and patents from Digicash

Founded by Chaum, went into Chapter 11 in 1998

Has not picked up as fast as hoped

- Credit card companies are putting up fight and transactions are becoming more efficient
- **Government is afraid of abuse**

Currently mostly used for Gift Certificates, but also used by Deutsche Bank in Europe.

15-853

Page 45

The Perfect Crime

- Kidnapper takes hostage
- Ransom demand is a series of blinded coins
- Bank forced to sign the coins to pay ransom, and publish the coins in the newspaper (they're just strings)
- Only the kidnapper can unblind the coins (only he knows the blinding factor)
- Kidnapper can now use the coins and is completely anonymous

15-853

Page 46

Chaum's protocol for offline anonymous cash

How do we prevent double payment without bank intervention?

Idea:

- If used properly, Alice stays anonymous
- If Alice spends a coin twice, she is revealed
- If Merchant remits twice, this is detected and Alice remains anonymous
- Must be secure against Alice and Merchant colluding
- Must be secure against one framing the other.

An amazing protocol

15-853

Page 47

Chaum's protocol: money orders

u = Alice's account number (identifies her)

r_0, r_1, \dots, r_{n-1} = n random numbers

(u_i, ur_i) = a secret split of u using r_i ($0 \leq i < n$)
e.g. using $(r_i, r_i \text{ xor } u)$

vl_i = a bit commitment of all bits of u_i

vr_i = a bit commitment of all bits of ur_i

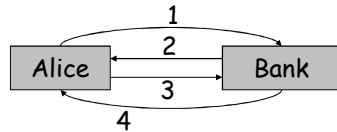
Money order:

- Amount
- Unique ID
- $(vl_0, vr_0), (vl_1, vr_1), \dots, (vl_{n-1}, vr_{n-1})$

15-853

Page 48

Chaum's protocol: Minting



1. Two blinded money orders and Alice's account #
2. A request to unblind and prove all bit commitments for one of the two orders (chosen at random)
3. The blinding factor and proof of commitment for that order
4. Assuming step 3. passes, the other blinded order signed

15-853

Page 49

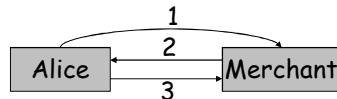
Can Alice Cheat the Bank?

1. With probability $\frac{1}{2}$, Alice can get the bank to sign coins that don't match her account #.
2. How would you fix this?
 1. If caught, impose a large fine (Alice must sign request for money order, so we know its her)
 2. Break up money order into n smaller coins, each worth $1/n$ of the order. Run the protocol with each coin. If Alice cheats on K coins, probability of bank not detecting this is $(1/2)^K$.

15-853

Page 50

Chaum's protocol: Spending



1. The signed money order C (unblinded)
 2. A random bit vector B of length n
 3. For each i :
 - if $B_i = 0$ return bit values for u_i
 - else return bit values for u_i
 Include all "proofs" that the u_i or $u_i r_i$ match v_i or $v_i r_i$
- Now the merchant checks that the money order is properly signed by the bank, and that the u_i or $u_i r_i$ match the v_i or $v_i r_i$

15-853

Page 51

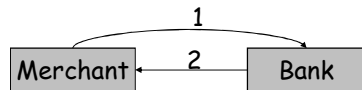
Can Alice Anonymously Spend a Coin Twice?

- Two merchants pick two bit vectors B, B'
- Alice provided to merchant #1:
 - u_i for all i such that $B_i = 0$
 - $u_i r_i$ for all i such that $B_i = 1$
- Alice provided to merchant #2:
 - u_i for all i such that $B'_i = 0$
 - $u_i r_i$ for all i such that $B'_i = 1$
- Merchants provide all this info to the bank. If there exists i such that $B_i \neq B'_i$, then bank gets $(u_i, u_i r_i)$ and identifies Alice.
- $\Pr[\text{Alice succeeds}] = (1/2)^n$

15-853

Page 52

Chaum's protocol: Returning



1. The signed money order
The vector B along with the values of u_i or ur_i that it received from Alice.
 2. An OK, or fail
- If fail, i.e., already returned:
1. If B matches previous order, the Merchant is guilty
 2. Otherwise Alice is guilty and can be identified since for some i (where B s don't match) the bank will have (u_i, ur_i) , which reveals her secret u (her identity).

15-853

Page 53

Miller-Rabin Primality Test

Write $n = 2^k m + 1$, m odd
Pick random integer x , $1 < x < n$
Let $y = x^m \bmod n$
If $y = 1 \bmod n$ then return "evidence for primality"
Else for $i = 0$ to $k-1$
 if $y = -1 \bmod n$ return "evidence for primality"
 else $y = y^2 \bmod n$
Return "certainly composite"

15-853

Page 54