CS6431: Security and Privacy Technologies

$c = m^e \mod n$

Cryptocurrencies

Instructor: Ari Juels
(c) 2015 Ari Juels
Historical backdrop

One of the earliest proposed uses of digital signatures (RSA) was to create virtual currency (in Ireland)

- Idea: A bank creates coins consisting of digital signatures
- Simplified version...

\[
\text{Virtual } \$$ = \text{Sig}(SK, \text{“I’m dollar #123”})
\]

Everyone can check: \( \text{Ver}(PK, \text{Virtual } \$$) \)

\( (SK, PK) \)

Virtual $\rightarrow$ Alice $\rightarrow$ Virtual $\rightarrow$ Bank $\rightarrow$ Virtual $\rightarrow$ Bob
E-cash

- Bank can also record spent serial numbers
- Bob can verify validity of coin online

\[
\text{Virtual } \$ = \text{Sig}(SK, \text{“I’m dollar #123”})
\]

\[
\text{Ver}(PK, \text{Virtual } \$)
\]

Everyone can check:

Alice

Bob
E-cash

- Privacy problem?

\[ \text{Virtual } $ = \text{Sig}(SK, \text{“I’m dollar #123”}) \]

Everyone can check: \[ \text{Ver}(PK, \text{Virtual } $) \]

\[ (SK, PK) \]

Virtual $ \rightarrow$ Virtual $ \rightarrow$ Virtual $ \rightarrow$ Virtual $
Blind digital signatures (Chaum, 1982)

• Ideas:
  • Alice chooses a serial number \( z \) for virtual $;
  • Bank digitally signs \( z \) without seeing it

• RSA setup:
  • Public key: \((N, e)\); private key \((N, d)\)
  • Full-domain hash: \( H: \{0,1\}^* \rightarrow \mathbb{Z}_N \)
RSA blind signature

serial num. $z \leftarrow \$ 
$c = H(z)$ 
$r \leftarrow \$ Z_N$

t = r^e c$

$s' = t^d$

$s' = r c^d$

$s = s' / r = c^d$

RSA public key $(N, e)$; private key: $(N, d)$
Operations mod $N$
What if we want to go offline?

• Suppose Mallory double-spends in non-anonymous scheme?

\[(SK, PK)\]

Virtual $ #123

Virtual $ #123

Virtual $ #123

Mallory
What if we want to go offline?

- Suppose Mallory double-spends in blinded scheme?

\( (SK, PK) \)

Virtual $ #123

Mallory

Virtual $ #123

Virtual $ #123

Virtual $ #123
Conditional anonymity

Intuition:

- Identity encoded as value $b$
- Receiver challenges coin spender to reveal point $f(a)$ when spending
- Double-spending: two points uniquely specify line!

$$ f(x) = ax + b $$
More history

- In 1982, “blind” digital signatures introduced by David Chaum
  - (Yes, the same Chaum who invented mix networks ➔ Tor)
  - Many research problems
    - Identity escrow
    - Turned into Digicash company [1990-1998]
- Researchers published hundreds of papers on virtual currency for decades after
  - Financial Cryptography conference [1997-]
  - PayWord and MicroMint (Rivest and Shamir) [2001]
    - MicroMint used proof of work to create coins
    - PayWord gave rise to Peppercoin company [2001-7]
- 840 Google Scholar citations (more than Bitcoin paper as of 16 Nov. 2015!)
- But no one used virtual currency until…

More history
Bitcoin

• Created by “Satoshi Nakamoto”
  • Paper “Bitcoin: A peer-to-peer electronic cash system” [2008]
  • Source code [2009]
• As of today, $6+ Billion market capitalization
  • But who is Nakamoto?

Bitcoins’s face

Maybe not.
Bitcoin

- Created in 2008 by “Satoshi Nakamoto”
  - Paper “Bitcoin: A peer-to-peer electronic cash system”
  - Source code
- As of today, $6+ Billion market capitalization
- But who is Nakamoto?

Another theory…
• Created in 2008 by “Satoshi Nakamoto”
  • Paper “Bitcoin: A peer-to-peer electronic cash system”
  • Source code

• As of today, $6+ Billion market capitalization

• But who is Nakamoto?
Bitcoin design—from basic principles
Key property #1: Bitcoin is pseudonymous

- What does this mean?
- Each entity $X$ has an (ECDSA) key pair $(SK_x, PK_x)$
- No association between $X$ and real-world identity
Digital signatures are used in Bitcoin

Bitcoin uses ECDSA

- “Elliptic-Curve Digital Signature Algorithm”
- Concretely, uses secp256k1 (slightly nonstandard) curve
  - Private key $SK$ is 256 bits; (uncompressed) public key $PK$ is 512 bits
Could build naïve system...

- Idea: Coins *and* transactions, i.e., flow of money, can be authenticated—neither is forgeable.
- Thanks to public-key crypto, everyone can verify all coins and transactions (if public keys are distributed throughout system).

\[
\text{Sig}(SK_x, \text{“Virtual $#123 \text{ to } Y”})
\]

\[
(SK_x, PK_x) \quad \Rightarrow \quad (SK_y, PK_y)
\]
But there’s a problem…

Double spending!!

Mallory

\( (SK_M, PK_M) \)

\[ \text{Sig}(SK_M, \text{“Virtual $ #123 to Y”}) \]

\[ \text{Sig}(SK_M, \text{“Virtual $ #123 to } X”) \]
The double spending problem

• Suppose Mallory gives the same dollar (dollar #123) to both Alice and Bob.

• The good news:
  • She can be caught after the fact.
  • Alice and Bob have (cryptographic) proof that Mallory cheated: Mallory’s signatures!

• The bad news:
  • Either Alice or Bob is out a dollar.
  • Who’s going to prosecute Mallory?
  • E.g., suppose Mallory is halfway around the world?
Idea: Bank maintains a ledger

\[
\text{Sig}(SK_x, \text{ “Virtual $#123 to } PK_y”) \]

Bob checks Sig and ledger

\((SK_x, PK_x)\)  \(\rightarrow\)  \(\rightarrow\)  \(\rightarrow\)  \(\text{Sig}(SK_x, \text{ “Virtual $#123 to } PK_y”)\)  \(\rightarrow\)  \(\rightarrow\)  \(\rightarrow\)  \(\text{Sig}(SK_Y, PK_Y)\)
Ledger

- Ledger is up-to-date record of all transactions.
- Bob now checks the ledger to be sure that Virtual $ #123 hasn’t been spent.
- Double-spending is now prevented!
But there’s still a problem…

- You have to trust the Bank!
- Problems:
  - What if the Bank claims not to have received a transaction?
    - i.e., doesn’t put it in ledger
  - What if the Bank confiscates money?
  - Who’s going to create money? The Bank?
  - What if the Bank devalues money?
Key property #2: Bitcoin is decentralized

- No Bank!
- Ledger is agreed upon and distributed among many entities
- Called the blockchain in Bitcoin
- The key innovation in Bitcoin over older virtual currencies
Blockchain

- Record of *every* transaction in Bitcoin system
- Maintained as append-only data structure
- New block added every 10 minutes (on average)
- Each block contains a bundle of latest transactions.
  - E.g., $\text{SIG}_{\text{PKA}}[\text{“Alice sends 0.4 BTC to Bob”}]$
  - (Actually, there’s a scripting language, but we’ll gloss over it...)

```
height = 0
height = 1
height = 2
height = 3
```
Blockchain

- Because full chain is a complete ledger/history of all transactions...
- Computing over the full block chain reveals the state/ownership of all BTC
- No explicit “account balances”
- Structured in terms of transactions

[Figure source: http://www.righto.com/2014/02/bitcoins-hard-way-using-raw-bitcoin.html; Hi, Ken!]
But how is a block validated?

• I.e., how does the system decide what transactions go into the next block?

• Ideal for P2P system: All clients in the world vote on the correct block chain.

• But it’s hard to ensure one vote per machine.

  • E.g., there’s the problem of “Sybil” attacks: How to prevent one user from creating multiple identities?

• So "voting" (cleverly) in Bitcoin takes the form of hash power.

  • I.e., one vote per CPU (roughly speaking)
But how is a block validated?

- Communal, computationally-intensive process called *mining*.
  - Together, mining community defines blockchain

Intuition:
  - All miners collectively search for hard-to-compute “signature” on new block
  - Solution proves w.h.p. that result is communal effort
  - Attacker with little computing power unlikely to mine block
Block N+2

Block N+1

Block N

Precise mining problem: Find a ticket that yields hash image with value less than target Z
Block mining

Block_{N+2} 

SHA-256^2(Block_{N+1}, X_{N+1}, ticket_{N+1}) = 0x00000000000000001d7a1...

Block_{N+1} 

SHA-256^2(Block_{N}, X_{N}, ticket_{N}) = 0x0000000000000000c67aa...

X_{N} = (software version, hash (Merkle-tree root) of new transactions, and current time)
This problem requires a massive amount of hash power

- The mining puzzle is called a *Proof of Work (PoW)*
- In Random Oracle Model for SHA-256, expected (double) hashes to mine a block is...
  - $2^{256} / Z$
  - $= \text{Bitcoin “Difficulty”} \times 2^{32}$
- Difficulty adjusted every 2016 blocks to achieve 10-minute block mining epoch

$$\text{SHA-256}^2(\text{Block}_{N+1}, X_{N+1}, \text{ticket}_{N+1}) \leq Z$$
This problem requires a massive amount of hash power

• Today (Nov 2015), expected number of (double) hashes to mine a block is roughly $340,000,000,000,000,000,000$.

• Whole Bitcoin network is running at about $520,000,000$ GH/s!
  • Roughly $2^{80}$ effort per week
  • (Estimated effort to factor 1024-bit RSA!)

• Not easily duplicated, so hard for attacker to seize control of network

SHA-256^2(Block_{N+1}, X_{N+1}, ticket_{N+1}) \leq Z
What’s the incentive for miners to mine?

- Key idea: Bitcoin is a lottery.
- Every miner tries tickets until a “winning” one is found.
- The prize for the winner: Bitcoins!
  - Special transaction in block assigns BTC to winner
  - Originally, 50 BTC; today, 25 BTC ($9,000+)
  - Winner also gets transaction fees
- 21 million BTC will be produced over the lifetime of the system.

SHA-256^2(\text{Block}_{N+1}, X_{N+1}, \text{ticket}_{N+1}) \leq Z

Courtesy: Brian Warner

Nov. 2015
height \approx 383k
What’s the incentive for miners?

- In principle, Bitcoin is democratic
- *Anyone* can mine.
- Reward is proportional to computational investment.
- But…
How do miners mine?

- In the early days, people just used their PCs.
- ASIC (Application-Specific Integrated Circuit) hardware is much more cost-effective.
- Professionals buy and replace ASICs frequently.

**Major Update (September 10, 2014):** Speed increase; 6TH/s Yukon is now $3,920.00!! (Best on the Market for now)

**The SP31 Yukon Power Miner**

The introduction of the SP31 Yukon powerful miner is good news to the bitcoin markets. The essential 5.5 TH/s mining machine focuses on the affairs of traders, it has an amazing hash power and consumes relatively very low power. It has been understood that the hashing power of SP31 has four times the power of SP10 and relatively twice the hash to power ratio.
Mining pools

Hashrate Distribution

An estimation of hashrate distribution amongst the largest mining pools

The graph below shows the market share of the most popular bitcoin mining pools. It should only be used as a rough estimate and for various reasons will not be 100% accurate. A large portion of Unknown blocks does not mean an attack on the network, it simply means we have been unable to determine the origin.

24 hours - 48 hours - 4 Days

Researchers from Cornell University say that on multiple occasions, a single mining pool repeatedly contributed more than 51 percent of Bitcoin's total cryptographic hashing output for spans as long as 12 hours. The contributor was GHash, which bills itself as the "#1 Crypto & Bitcoin Mining Pool." During
GINI Index

Estimate for Bitcoin: 88
Other parts of Bitcoin
Mining blocks isn’t enough

- What else is needed to make a working monetary system?
  - Broadcasting transactions and blocks
  - Storing ledger / blockchain
  - Enabling users to spend and receive money
Some node types in Bitcoin network

**Reference Client (Bitcoin Core)**

Contains a Wallet, Miner, full Blockchain database, and Network routing node on the bitcoin P2P network.

**Full Block Chain Node**

Contains a full Blockchain database, and Network routing node on the bitcoin P2P network.

**Solo Miner**

Contains a mining function with a full copy of the blockchain and a bitcoin P2P network routing node.

**Lightweight (SPV) wallet**

Contains a Wallet and a Network node on the bitcoin P2P protocol, without a blockchain.

[Source: http://chimera.labs.oreilly.com/books/1234000001802/ch06.html]
Routing functionality

- Transactions and blocks are broadcast to entire network of nodes
- Rebroadcast protocol
  - Each node transmits to 8 other (randomly selected) nodes
  - TCP on port 8333
Storage

• Full nodes:
  • Store entire blockchain
  • Enforce consensus rules, ensuring blocks in blockchain adhere to
    • 25 BTC reward
    • Correct signatures on transactions
    • BTC not double-spent
    • Etc., etc.
Full node distribution

BITNODES
Bitnodes is currently being developed to estimate the size of the Bitcoin network by finding all the reachable nodes in the network.

GLOBAL BITCOIN NODES DISTRIBUTION
Reachable nodes as of Mon Nov 10 2014

6704 nodes
24-hour charts »

Top 10 countries with their respective number of reachable nodes are as follow.

<table>
<thead>
<tr>
<th>RANK</th>
<th>COUNTRY</th>
<th>NODES</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>United States</td>
<td>2956</td>
</tr>
<tr>
<td>2</td>
<td>Germany</td>
<td>506</td>
</tr>
<tr>
<td>3</td>
<td>Canada</td>
<td>421</td>
</tr>
<tr>
<td>4</td>
<td>France</td>
<td>379</td>
</tr>
<tr>
<td>5</td>
<td>United Kingdom</td>
<td>371</td>
</tr>
<tr>
<td>6</td>
<td>Netherlands</td>
<td>278</td>
</tr>
<tr>
<td>7</td>
<td>Russian Federation</td>
<td>188</td>
</tr>
<tr>
<td>8</td>
<td>China</td>
<td>178</td>
</tr>
<tr>
<td>9</td>
<td>Australia</td>
<td>129</td>
</tr>
<tr>
<td>10</td>
<td>Sweden</td>
<td>118</td>
</tr>
</tbody>
</table>

More (87) »

Map shows concentration of reachable Bitcoin nodes found in countries around the world.
Bitcoin wallets

• You don't need to mine or run full node to use Bitcoin.
• Wallets are applications that permit easy management of a Bitcoins.
• What’s going on under the hood?
Bitcoin wallets: Under the hood

- Remember: identity associated with ECDSA digital signature key pair
  - $SK$ used to sign / authorize transactions.
  - $PK$ used to identify users and verify transactions.
- Bitcoin wallet stores, protects, and allows use of $SK$ to make transactions.

Credit: Ken Shirriff
Brain wallets

- You can generate SK from a password

Your Bitcoin are then completely portable.
Brain wallets

- Unfortunately, human brains are poor password stores...
- Cracking brain wallets at one point rumored more profitable than mining...

Finders keepers? I found an address with 50 BTC via brain wallet!
January 18, 2014, 04:58:04 PM

I was playing around with the brain wallet and checking the addresses with blockchain. I found a wallet with a balance of 50 BTC! The coins were put in the wallet in 2011 and there hasn't been any activity since. I don't want to steal someone's coins but if they are "lost" I don't want to have them just sitting there. It's a lot of money! I was thinking of sending a small amount into the wallet with a message letting the person know the situation. If nothing happens after a while I guess it's "Finders Keepers, Losers Weepers." What's the right thing to do in this situation?
Bitcoin use today
Online sites

1. **Get Bitcoin**

   There are several ways to get Bitcoins, but the easiest is to exchange them for currency at your bank or a Bitcoin exchange. You can also buy Bitcoins from friends, accept them as payment for goods or services, or generate new Bitcoins through a process called “mining.”

   [Sign Up at Coinbase.com](#)

2. **Shop Overstock.com**

   You can now pay for all your favorite products on Overstock.com using Bitcoins! As the first major retailer to accept Bitcoins, Overstock.com is expanding the possibilities of Bitcoin purchases by offering thousands of products to the Bitcoin community.
Yet another reason to visit Cornell Tech…
Bitcoin exchanges

It seems everyone loves Bitcoin!

Including your dog…

Laura Saggers releases world’s first Bitcoin love song

Business Review

Bitcoin Financial Report

Chart taken from BitcoinCharts.com

bitcoins pet tag
$21.95

@bitcoin
ACCEPTED HERE
...except environmentalists!
...except environmentalists

• $250++ million in computing hardware invested in Bitcoin ecosystem
• 433+ MW ([http://realtimebitcoin.info/](http://realtimebitcoin.info/))
  • More than largest solar power station in U.S. at Kramer Junction (354MW)

and growing...
Tor + Bitcoin = End-to-end anonymity for commercial transactions
...and law enforcement

Bitcoin has stimulated

- Money laundering
- Theft of Bitcoin wallets
- Illicit marketplaces (Silk Road)
- Rogue mining
  - E.g., ZeroAccess botnet
- Ransomware
Not truly anonymous

• Recall Bitcoin is *pseudonymous*, i.e., traceable on per-identity basis
• E.g., suppose you’re Satoshi Nakamoto and you want to spend your 1,624,500 BTC ($500+ million) anonymously…
• Thus NSA conspiracy theory…

But...

Zerocash

Zerocash is a new protocol that provides a privacy-preserving version of Bitcoin (or a similar currency).

Zerocash fixes an inherent weakness of Bitcoin: every user's payment history is recorded in public view on the blockchain, and is thus readily available to anyone. While there are techniques to obfuscate this information, they are problematic and ineffective. Instead, in Zerocash, users may pay one another directly, via payment transactions that reveal neither the origin, destination, or amount of the payment. This is a marked improvement compared to Bitcoin (and similar decentralized digital currencies), where every payment's information is made public for the whole world to see.

Zerocash improves on an earlier protocol, Zerocoin, developed by some of the same authors, both in functionality (Zerocoin only hides a payment's origin, but not its destination or amount) and in efficiency (Zerocash transactions are less than 1KB and take less than 6ms to verify).
New York one of the first states to regulate Bitcoin

Know-your-customer (KYC) at odds with pseudonymity / anonymity!
Long-term problems

• Scaling!
• Blocks are at most 1MB in size
  • Transaction about 500B on average
  • At most 2000 transactions per block
  • 3 transactions / sec. throughput
• Should we:
  • Increase the block size or
  • Increase the mining rate or
  • Do something else?
• Big controversy!
Long-term problems

• Limited bandwidth
  • Recent measurement (2014-5) indicates that the 10%, median, and 90% block propagation times are 0.8 seconds, 8.7 seconds, and 79 seconds respectively
  • 540k block size

(a) Network throughput vs. block size.
Long-term problems

- Lightweight client (SPV, Spectrum, Pool) sufficient for wallet / mining
- No monetary incentive to run full node
  - 100,000 miners, but only 5000+ full nodes
  - Bitnodes actually offering reward for full nodes!

**BITNODES INCENTIVE PROGRAM**

Bitnodes incentive Program is an experimental incentive program to allow reachable nodes with Bitcoin address set to receive weekly incentive paid in bitcoins. The program ends by Thu Dec 31 2015 23:59:59 GMT+0000 (GMT) or when the network sustains 10000 or more reachable nodes for 24 hours, whichever comes first.

Bitcoin address for a reachable node can be set using the Bitnodes API. The weekly incentive will be paid in bitcoins to the Bitcoin address of a node selected randomly from a pool with at least 100 eligible nodes. A node is considered eligible if it has a verified Bitcoin address set and its Peer Index (PIX) value is greater or equal to 8.0.

The weekly incentive varies according to the number of reachable nodes as per the schedule below. The amount of bitcoins will be calculated using CoinDesk BPI and will be paid out from my personal funds throughout the duration of the program.

<table>
<thead>
<tr>
<th>REACHABLE NODES</th>
<th>WEEKLY INCENTIVE</th>
</tr>
</thead>
<tbody>
<tr>
<td>&gt;= 5000</td>
<td>USD 10.00</td>
</tr>
<tr>
<td>&gt;= 7000</td>
<td>USD 20.00</td>
</tr>
<tr>
<td>&gt;= 9000</td>
<td>USD 30.00</td>
</tr>
</tbody>
</table>
Full node functionality

Bitnodes is currently being developed to estimate the size of the Bitcoin network by finding all the reachable nodes in the network.

6704 nodes
24-hour charts »
Top 10 countries with their respective number of reachable nodes as of僧心情 that 21:29:39 GMT-0500 (EST).

5282 NODES
24-hour charts »
Top 10 countries with their respective number of reachable nodes are as follow.

<table>
<thead>
<tr>
<th>RANK</th>
<th>COUNTRY</th>
<th>NODES</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>United States</td>
<td>1924</td>
</tr>
<tr>
<td>2</td>
<td>Germany</td>
<td>619</td>
</tr>
<tr>
<td>3</td>
<td>France</td>
<td>389</td>
</tr>
<tr>
<td>4</td>
<td>Netherlands</td>
<td>280</td>
</tr>
<tr>
<td>5</td>
<td>United Kingdom</td>
<td>252</td>
</tr>
<tr>
<td>6</td>
<td>Canada</td>
<td>235</td>
</tr>
<tr>
<td>7</td>
<td>Russian Federation</td>
<td>172</td>
</tr>
<tr>
<td>8</td>
<td>Sweden</td>
<td>116</td>
</tr>
<tr>
<td>9</td>
<td>Australia</td>
<td>84</td>
</tr>
<tr>
<td>10</td>
<td>China</td>
<td>79</td>
</tr>
</tbody>
</table>
What do people / media mean by the **blockchain**?

- Nebulous term…
- Generally refers to *ledger*
- Distributed, robust, publicly visible piece of memory
- Good for things other than money!
  - Timestamping documents
  - Audit
  - Etc., etc.
What’s a (decentralized) smart contract?

- Executable object on blockchain
- Scripted in Turing-complete language
  - Bitcoin has highly restricted language
- Code defines contract, e.g.,
  - Financial instrument
    - If GOOG rises to $1,000 by 30 June 2015, assign 10 shares from Alice to Bob and pay Alice $10,000
  - Szabo (1997): Smart contract reassigns physical access to your car from you to your bank if you don’t make a payment
- Autonomous: *Enforced / executed by network*
What’s a (decentralized) smart contract?

• Executable object on blockchain
• Scripted in Turing-complete language
  • Bitcoin has highly restricted language
• Code defines contract, e.g.,
  • Financial instrument
    • If GOOG rises to $1,000 by 30 June 2015, assign 10 shares from Alice to Bob and pay Alice $10,000
  • Szabo (1997): Smart contract reassigns physical access to your car from you to your bank if you don’t make a payment
• Autonomous: \textit{Enforced by network}
What’s a (decentralized) smart contract?

- Executable object on blockchain
- Scripted in Turing-complete language
- Bitcoin has high restricted smart-contract language
- Contract defined by code, e.g.,
  - Financial instrument
    - If GOOG rises to $1,000 by 30 June 2015, assign 10 shares from Alice to Bob and pay Alice $10,000
- Szabo (1997): Smart contract reassigns physical access to your car from you to your bank if you don’t make a payment
- Autonomous: Enforced by network

**Intuition:** Smart contract simulates trusted third party with public state.
Simple smart contract: Lottery

**Init:**
- `T`end := 30 June 2015, $ticket := 1, pool := {}
- pot := 0

**TicketPurchase:**
- On receiving $amt from some party P:
  - Assert $amt = $ticket, balance[P] ≥ $amt
  - pot := pot + $ticket
  - pool := pool ∪ P

**Timer:**
- If T > Tend then
  - W ∈ R pool
Authenticated data feeds

- Remember simple smart contract example:
  - If GOOG rises to $1,000 by 30 June 2015, assign 10 shares from Alice to Bob and pay Alice $10,000

- Problem: Contract needs to know the price of GOOG!

- Community solution: ecosystem of authenticated data feeds

\[
\text{Sig}(SK_{\text{authority}}, T, \text{price}, "\text{GOOG"})
\]
Smart contract systems will have many data feeds…

…digitally signed by (trustworthy) sources.
Research problems
Many research problems / disciplinary perspectives

- Balancing anonymity / crime tension
  - Resuscitate de-anonymization?
  - Vote to ban bad transactions / smart contracts?
- Scaling approaches
  - BitcoinNG
  - Consortium cryptocurrencies
- Programming languages and smart contracts
- Etc., etc…