Odds and Ends
• Blockchain/cryptocurrency
• More on anonymity
• More on human-centered security
Blockchain,
bitcoin
Cryptocurrency goals

• (approximate)

• Decentralized

• Anonymous — analogize to cash

• Reliable
What is a blockchain?

- Essentially, a linked list with hashes as pointers
- Goal: Cannot add or delete blocks from the middle
  - Why?
- Content can be anything
BTC + Blockchain

- What content? Payment ledger
  - List of who paid who

- Account ID (address) = public key
  - To pay, sign with secret key

- Transactions aggregated into validated blocks
  - Check signature, check valid $, etc.

- If simultaneous blocks, longest chain wins
Potential problems

- Refuse to relay transactions (peer-peer)
- Exclude transactions from blocks
- Spend non-existent BTC
- Try to orphan blocks
  - Double-spend attack
• Refuse to relay transactions (peer-to-peer)

• Exclude transactions from blocks
  • OK as long as network is large, decentralized!

• Spend non-existent BTC
  • Not possible b/c ledger from beginning of time
Double-spend attack

- A writes a block paying B with BTC#1234
- A also writes a block paying BTC#1234 to A’
  - And a follow-up block to make a longer chain
- Need to make it hard to do this!
Proof of work

• Make it hard to create a valid new block!
  • So no one can monopolize it, create long chains at once, etc.

• Hash of the block contents is *salted*

• Successful mine: choose salt s.t. hash < target
  • Target = “network difficulty”
  • Shrinks over time so blocks remain slow

• Mining “reward” creates new BTC (for now)
Bitcoin + Anonymity

- Not anonymous, pseudonymous!
  - When your address is ID’d you are ID’d
  - Multiple recipient addresses -> one payment
- Transition to real money is hard to do anonymously
- Newer alternatives (e.g., zerocash)
Back to Tor
Recall: The Onion Router

- End-to-end path = a circuit
  - Default = 3-hop circuits
  - Download a big list of available peers
- Each middle node only knows its links
Encryption layers

\[ E(k_{M3}, M4 \ || \ E(k_{M4}, R \ || \ E(k_R, m))) \]

\[ E(k_{M4}, R \ || \ E(k_R, m)) \]

\[ E(k_R, m) \]
Something is still missing …

- We have disguised senders, what about receivers?
- Goal: Run service X on host D
  - Without anyone knowing D runs it
    - *hidden service*
  - (aka, dark web)
Hidden services

• Bob creates his service
  • Set up circuits to *introduction points*
  • Create a directory listing that maps X to points

• Alice wants to connect
  • Set up circuit to *rendezvous point* R
  • Associate with unique token I
  • Set up circuit to one of the intro points
  • Send message: Please forward R, I to X
Step 4: Alice writes a message to Bob (encrypted to PK) listing the rendezvous point and a one-time secret, and asks an introduction point to deliver it to Bob.
Hidden services (2)

- Connection via R
  - Bob sends message containing I to R
  - R links the two circuits together (forwarding)
  - Alice and Bob can now talk anonymously
Step 5: Bob connects to the Alice's rendezvous point and provides her one-time secret.
Who knows what?

• Only Bob knows he runs service X

• Intro point knows someone accessed X, but not who

• R knows someone accessed a hidden service, but not who or what

• Alice knows she accessed X, but not who/where X is
Potential Tor attacks

• Insert malicious relays into the network
  • Or compromise legitimate ones
  • Generally need multiple to be useful

• DOS on trustworthy routers
  • Drive traffic toward your relay

• DOS more generally
  • Force relay to do expensive crypto a lot
More Tor problems

- Exit nodes can be blamed for abusive actions
  - Limits desire to be an exit node
  - Monitor exit nodes for traffic analysis
- Option/configuration issues / fingerprinting
(Browser) Fingerprinting vs. Anonymity
What is fingerprinting?

• Using browser characteristics (fonts, screen dimensions, clock skew etc.) to uniquely ID

• *Does not require* client-side storage
  • Unlike traditional cookies
  • Works fine even in private browsing mode

• In 2010, 83% (of almost 500k users) were unique!
  • [panopticlick.eff.org](https://panopticlick.eff.org)

Eckersley, PETS 2010
Legitimate uses

- Preventing DOS
- Preventing fraud or account hijacking
- Identify content scrapers

- ... but also tracking with no consent, no opt out

Font probing

- Using JavaScript, load fonts and measure
  - In 2013, 13 scripts on 404 sites in Alexa top 100k

- Using Flash, enumerate directly

- Mainly anti-fraud and analytics companies
  - Ad campaigns, newspaper paywalls
  - But also anonymizer.com, Coinbase

Acar, CCS 2013
Canvas fingerprinting

- Draw text on Canvas API
  - Varies w/ OS, font library, graphics card/driver, browser, rasterization, physical display …
  - Retrieve via `dataURL` – binary pixel data, then hash
- Like font probing, no local storage
- Estimate: No more than 1/1000 overlaps

Mowery + Shacham, 2012
Acar, CCS 2014
Canvas fingerprinting in the wild

• Survey of Alexa top 100k sites: home pages
  • See paper for interesting detection details

• More than 5.5% actively using
  • Vast majority via addthis.com

• Additional techniques
  • Draw in 2 different colors
  • Use fake font name to get default font
  • Cwm fjordbank glyphs vext quiz, 😊

<table>
<thead>
<tr>
<th>Alexa range</th>
<th>% using</th>
</tr>
</thead>
<tbody>
<tr>
<td>[1, 1k)</td>
<td>1.8</td>
</tr>
<tr>
<td>[1k, 10k)</td>
<td>4.9</td>
</tr>
<tr>
<td>[10k, 100k]</td>
<td>5.7</td>
</tr>
</tbody>
</table>

Acar, CCS 2014
Cookie abuse

• Cookie syncing: 3rd-party domains sharing IDs
  • e.g., via HTTP referer

• Evercookies: respawn cleared cookies via flash, HTML 5, canvas cache, etc. etc.
Cross-device targeting

- Explicit: Same account on multiple devices
- Implicit: Related searches from same geo. location
- Sneaky: Generate/listen for high-pitched sounds
Countermeasures & mitigations

• Canvas: ask on all data reads?
  • Can’t disable entirely without breaking functionality

• Evercookies
  • Clear lots of storage locations
  • Browser mechanisms are not straightforward
  • e.g., Flash across browsers

• Cookie syncing:
  • 3rd party cookie blocking
  • But only from fresh state!
Countermeasures & mitigations

• Tor browser
  • Fixed settings to prevent differentiation
  • Cap on font enumeration
  • Return empty object from canvas reads
  • Clear huge list of storage caches

• Assorted research tools
  • e.g., Firegloves extension

• Having Tor (or a research extension) is kind of unique to start with, though!