Anonymous Credentials: hCaptcha

Eli-Shaoul Khedouri
hCaptcha.com

Background

- Privacy-first humanity verification: CAPTCHA, risk scores, etc.

  Two user flows for anonymous credentials:

- General Verification of Humanity
  User → Website ← → hCaptcha [verifies challenge solved]

- Accessibility User
  User → hCaptcha [issues credentials] → Website ← → hCaptcha

Security issues: blinded redemption

- Risk data, rate limiting, etc. already associated with the IP
- If hidden on a per-redemption basis, either need to push this onto the application (bad) or create scoped, blinded identifiers

  One option:

  User → Website ← → hCaptcha Risk Score on IP; no PII sent

  User → Website ← → hCaptcha [Blinded IP, coarse risk score for IP]

- Accessibility

  User → hCaptcha [issues credentials] → Website ← → hCaptcha

Security issues: accessibility use case

- Can we provide a completely universal, privacy-preserving accessibility option to let some users skip the challenge?
- Idea: Use registration flow, risk evaluation at unblinded registration time determines whether to issue
- **However:** need to be able to segment by risk class in order to prevent gross abuse at redemption time
- Example: 16 fuzzy feature buckets encoded into “privacy byte”; can challenge redemptions when anomaly detection fires per-site

Other considerations: expiration etc.

- Running at large scale, need to design for distributed data; don't want to maintain global double spend cache
- Must be able to enforce at-will expiration, including bulk expiration of user segments or classes
- Require some coarse segmentation in blinded redemption case to enforce security guarantees

Privacy Pass: Practical Considerations

- Metadata (more than one bit) is required per-session to maintain security guarantees
- At-will expiration implies at-will key rotation, ideally rapid, if segmentation is done via key
- Must be able to deliver similar or better level of service quality, otherwise user or online service will opt out of privacy features
  - Native browser integration, consistent cross-browser policies