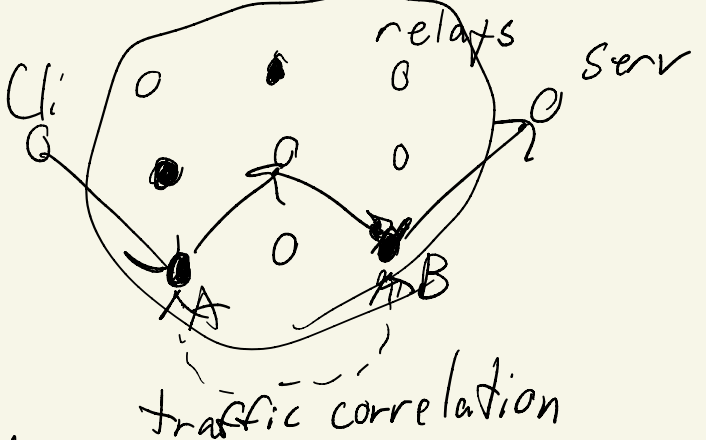
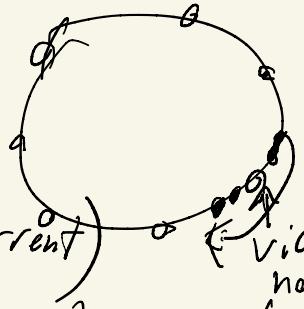
CS-234
Technologies for societal self-organization

Week 8

Sybil Attacks



- Tor network: Sybil Tor relays
- Distributed hash tables (IPFS)
- Consensus algorithms
- Fake/poisoned content (BitTorrent)

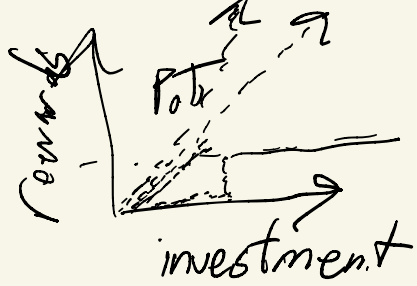


- Social bots: fake social media accounts
  - spread fake news (sow chaos, anger, affect opinion...)
    - Sybils amplify the believability/"reputation" of posts
      Likes/retweets etc + followers
      + algorithms for newsfeed selection
      Sybils manipulate the "network effect"
    - often hard to distinguish real/fake content
      often spreads more quickly than the truth
- Sockpuppets (Wikipedia) - get around restrictions
  - manipulate peer review
- Online voting/polling (Twitter polls)

Sybil defenses

- Detection / analysis (w/ help of user reports)
  - advantage: can benefit from AI/ML
  - disadvantage: false positives / negatives, attackers have AI/ML too
    - arms race winnable???
- Proof of Identity: authentication + authorization
  - phone numbers, unique HW identifiers, IP addresses
  - banking / financial KYC ID verification service
- Proof of Investment
  - CAPTCHAs: automated Turing tests: getting harder bc attacks getting better
  - Proof of work: Bitcoin - more work invested, more rewards
    - wastes energy!
  - Proof of stake: get existing coins stake (or lockup) for a time
    - eliminate energy waste
  - Proof of space / storage
- Proof of Personhood: "one person, one vote": "democratic"
  - natural

Proof of Personhood



- minimal investment, then plateau of rewards
- how to achieve?
  - strong identity - e.g. GoodDollar - face verified
    India Aadhaar - >1B people 2 eyes + 10 fingers each has FAR/FRR ≈ 1-in-100,000
    advantage: it works
    disadvantages: excluded, corruption, ..., privacy/surveillance
  - social trust networks (PGP "web of trust")
    Duniter, GPIs, UniqueID, Humanity DAO, ...
  - verification (mutual)
    Idena.io, Encointer,
  - pseudonym parties
    give one token per person @brynosaurus