Incentives in Security Protocols

Sarah Azouvi, Alexander Hicks, Steven J. Murdoch

Incentives are at the core of security economics

“In an ideal world, the *removal of perverse economic incentives* to create insecure systems would de-politicize most issues. Security engineering would then be a matter of rational risk management rather than risk dumping.”

Anderson – Why Information Security is Hard (2001)

Aligning incentives in security systems is easier said than done

- Incentives allow influencing of a system’s elements that cannot be tightly controlled through code or security protocols
- Can be non-economic (see security psychology field)
- Protocols do fail, but system can be made robust through proper alignment of incentives
- Fail-deadly alignment – person in position to prevent system failure will be harmed by the failure
- Fail-safe alignment – innocent parties will be protected from the consequences of system failure

EMV smart card payments rely on incentives for correct functioning

- Plenty of variants and plenty of failures
- Failures fall on a spectrum between designers’ being completely surprised and designer’s being well aware but believe incentives resolve the issue
- Cost of fraud falls on the party which led to insecure operation, to create a fail-safe overlay on top of a fragile protocol
- Only works if evidence is available to properly assign liability
- As a historical accident, communications follow contractual relationships so limits of EMV evidence are somewhat mitigated

Cryptocurrencies depend almost entirely on incentives for functioning

- Almost no contractual relationships, so consensus over what is the system state is through incentives around the formation of consensus
- Reasoning around incentives at the same stage that security protocols were in the 1980’s
- Failures result from synchrony assumptions and rational participants, in contrast to Byzantine failure
- Soft forks are a fail-safe mechanism, allowing older clients to function
- Chain split are fail-deadly, and do destroy value

Incentives can be non-economic, e.g. Tor

- Tor routes user traffic over volunteer relays (~6,000 today)
- No payment (though a handful are reimbursed for bandwidth)
- Fail-safe approach is to not rely on the security of any one relay
- Fail-deadly doesn't work when adversary can just come back again
- Monetary incentives have not yet been adopted, in part due to concern that this will discourage voluntary contribution

Discuss!

- How can we achieve incentive alignment in protocols, and make them a first-class object in protocol modelling (along with principals and keys)?
- How do we choose the right type of incentive?
  - Economic vs non-economic; internal vs external; explicit vs implicit; reward vs punishment
- How do we enable incentive enforcement?
  - Unambiguous, tamper resistant, interpretable evidence; trusted-third party; consensus
- What’s the right model for reasoning about incentives?
  - Nash equilibria; BAR model; Rational Cryptography