ON COMPUTATIONALLY EFFICIENT SYBIL DEFENSE

Diksha Gupta, Jared Saia and Maxwell Young
The Sybil Attack

John R. Douceur
Microsoft Research
johndo@microsoft.com

"One can have, some claim, as many electronic personas as one has time and energy to create." --- Judith S. Donath [12]

Abstract — Large-scale peer-to-peer systems face security threats from faulty or hostile remote computing elements. To resist these threats, many such systems employ redundancy. However, if a single faulty entity can present multiple identities, it can control a substantial fraction of the system, thereby undermining this redundancy. One approach to preventing these "Sybil attacks" is to have a trusted agency certify identities. This paper shows that, without a logically centralized entity, Sybil attacks are always possible except by collusion and unrealistic assumptions of coordination among entities.

If the local entity has no direct physical knowledge of remote entities, it perceives them only as informational abstractions that we call identities. The system must ensure that distinct identities refer to distinct entities; otherwise, when the local entity selects a subset of identities to redundantly perform a remote operation, it can be duped into selecting a single remote entity multiple times, thereby defeating the redundancy. We term the forging of multiple identities a Sybil attack [30] on the system.

It is tempting to envision a system in which established identities vouch for other identities, so that an entity can accept new identities by trusting the collective assurance of multiple (presumably independent) signatories, analogous to the PGP method of trust [37] for human entities. However, research show that, in the absence of a trusted authority (or unrealistic assumptions about the identity of an attacker), a Sybil attack can be deployed that appears to be the initial identity.
In an open system, an entity may claim multiple identities in order to gain unfair share of the network’s resources.

Specially challenging in the absence of a central certifying authority.
Specifically, attackers took advantage of the fact that anyone could operate on the network as a public Electrum peer. Attackers then launched what's called a Sybil attack that introduces compromised nodes into the network. The result of such an attack was that hundreds of thousands of computers have been compromised through the false security update and other means shown at the start of this article.
A technique widely popularized with the success of PoW-based cryptocurrencies.

Popular systems require solutions to be generated continuously in order to provide security guarantees.
WHY WE CARE ABOUT EFFICIENCY?
Energy Consumption by Country Chart

- Denmark: 33 TWh per year
- Belarus: 33.8 TWh per year
- Bulgaria: 34.9 TWh per year
- Bitcoin: 36.8 TWh per year
- Qatar: 39 TWh per year
- Hungary: 40.3 TWh per year
- New Zealand: 41.4 TWh per year

Source: https://digiconomist.net/bitcoin-energy-consumption

Jan, 2018
Bitcoin consumes more energy than Switzerland, according to new estimate

Country Ranking

- Switzerland: 58.46 TWh per year
- Czechia: 62.34 TWh per year
- Bitcoin: 64.15 TWh per year
- Austria: 64.60 TWh per year
- Colombia: 68.25 TWh per year
OUR MODEL
THE SYSTEM

Good IDs
- Follow the protocol.

Bad IDs
- Can deviate arbitrarily from the protocol, under the control of an adversary.
• **Random Oracle Assumption:** We have a function, h, such that \( h(x) \) is chosen uniformly at random on \((0, 1]\) the first time bit string x is input to h.

• **Computational Cost:** Number of times h is called.
Diffuse protocol:

- Every time a good ID sends a message, it is delivered to all other good IDs.

- Communication time is negligible compared to computation time.
ADVERSARY

- Has at most a small constant fraction of the computational power of the system.

- Knows our algorithms, not our random bits.
1st Protocol
Commensurate Computational - CCOM
**Theorem:** Let $T$ be the rate of adversarial spending and $J$ be the rate of join of good IDs. Then:

- There is always an honest majority.

- The rate of algorithmic spending is $O(J + T)$. 
INITIALLY…

Reduce the fraction of bad IDs to 1/3 using [1].

NEW ID JOINS

Here is my 1-difficult puzzle solution.
ONLY JOINS

1/3 fraction of new IDs join.
Purge

Solve a puzzle or be ejected.
Reduce the fraction of bad IDs to $\alpha < 1/3$. 
2ND PROTOCOL

Geometric Mean Computation - GMCOM
Theorem: Let $T$ be the rate of adversarial spending and $J$ be the rate of join of good IDs. Then:

- There is always an honest majority.
- The rate of algorithmic spending is $O(J + \sqrt{JT})$. 
OPTIMAL ENTRANCE COST

- Assume \( T = \mathcal{X} J \)
- Total entrance cost to good IDs: \( \theta(\mathcal{X} J_g \ell) \)
- Total purge cost to good IDs:
  \[
  |G| \leq |S| = \theta(J \ell)
  \]
- Spending rate of good IDs:
  \[
  \theta(J + \mathcal{X} J_g)
  \]
ENTRANCE COST FUNCTION

\[ e(k) = \frac{J}{J_g} \]

where

\( e(k) \) - entrance cost to the \( k^{th} \) ID joining in the current iteration

\( J \) - join rate in the current iteration

\( J_g \) - estimated join rate of good IDs
CANDIDATE JOIN RATE ESTIMATE

\[ J_g = \frac{\alpha |S|}{L} \]

where

\( L \) - Duration over which at least 2\( \alpha \) - fraction of current set of IDs joined

\( S \) - Set of IDs at the end of iteration
WHY IS THIS A GOOD ESTIMATE?

At the end of an iteration, fraction of bad IDs $\leq \alpha$
WHY IS THIS A GOOD ESTIMATE?

System differs by at most $2\alpha$ fraction
WHY IS THIS A GOOD ESTIMATE?

System differs by at most $2\alpha$ fraction of joins by good IDs $\geq \alpha$
**WHY IS THIS A GOOD ESTIMATE?**

Estimated join rate of good IDs \( \leq 2 \) (Actual join rate of good IDs) *

* Assuming that the join rate of good IDs does not change by more than a constant factor over successive iterations
Lower Bound: For any purge-based algorithm, there is an adversarial strategy ensuring the following for any iteration: The algorithmic spending rate is $\Omega(J + JT)$, where the algorithmic spending rate $T$, and $J$ are taken over the epoch.
$J_G = 2$ IDs/second  \hspace{1cm} 10K initial IDs and seconds  \hspace{1cm} \alpha = \frac{1}{14}$
WHAT NOW?
WHAT NOW?

- Committees are easy targets for an **adaptive adversary**
WHAT NOW?

- Committees are easy targets for an adaptive adversary

Evict good IDs at will, can equivocate and take over good IDs in the system.
Committees are easy targets for an adaptive adversary.

Evict good IDs at will, can equivocate and take over good IDs in the system.

Goal: GET RID OF COMMITTEE!
- Deciding when to purge

  Each ID decides for itself on seeing a 1/3 change in membership.

- Deciding on randomness for the purge:

  Solve an n-hard puzzle over a round and set the smallest solution as the random seed.
- What if the adversary solves hard puzzle?

  Force a purge on good IDs sooner.

  Adversary would have spent considerable computational effort, counterbalance the cost to good IDs.

- Our Guarantee:

  There is an honest majority in the view of all good IDs at all times.
WHERE TO NEXT?

- Rational + Adversary
  - Challenge - Define utilities capturing computational effort and player benefit from the system.
- Sybil resistant DHT + adaptive adversarial churn
  - Challenge - Preventing eclipse attack
Thank You!

Questions?
To be effective, these resource challenges must be issued to all identities simultaneously:

**Lemma 2:** If local entity $l$ accepts entities that are not validated simultaneously, then a single faulty entity $f$ can present an arbitrarily large number of distinct identities to entity $l$. 