The Sybil Attack on Peer-to-Peer Networks From the Attacker's Perspective

Andrew Rosen

Georgia State University

December 4, 2014

What Am I Going to Talk About?

- The Tor paper mentioned Sybil attacks [2], so you should have an idea of what they are.
- An attacker gives himself a greater presence in the network by pretending to have multiple identities.
- The Sybil attack is extremely well known, but there is little literature written from the attacker's perspective.

Distributed Hash Tables

I'm going to keep this to only the relevant info so we can get straight into the attack.

- Structured peer-to-peer (P2P) networks use distributed hash tables (DHT) as the organization backend.
- Nodes typically get an ID in the network by passing their IP address and port into a hash function.
  - This function is typically SHA1 [3], which will return a value from 0 to $2^{160} - 1$ (a 160-bit number).
  - The outputs of SHA1 are evenly distributed [1].

The goal of the Sybil Attack in A P2P network

See Whiteboard

- We want to inject a Sybil into as many of the regions between nodes as we can.
- The question I wanted to answer is what is the probability that a region can have a Sybil injected into it, given:
  - The network size n
  - The number of keys (IDs) available to the attacker (the number of identities they can fake).

Assumptions

- The attacker is limited in the number of identities they can fake.
  - To fake an identity, the attacker must be able to generate a valid IP/port combo he owns.
  - The attacker therefore has $num\_IP \cdot num\_ports$ IDs.
  - We'll set $num\_ports = 16383$, the number of ephemeral ports.
  - Storage cost is 320 KiB.
- I call the act of finding an ID by modulating your IP and port so you can inject a node *mashing*.
- In Mainline DHT, used by BitTorrent, you can choose your own ID at "random." The implications should be apparent.

Equations

The probability you can mash a region between two adjacent nodes in a size n network is:

$$P \approx \frac{1}{n} \cdot num\_ips \cdot num\_ports \tag{1}$$

An attacker can compromise a portion $P_{bad\_neighbor}$ of the network given by:

$$P_{bad\_neighbor} = \frac{num\_ips \cdot num\_ports}{num\_ips \cdot num\_ports + n - 1} \tag{2}$$

People like proofs, but I prefer to demonstrate with my simulation results so I can get onto questions.

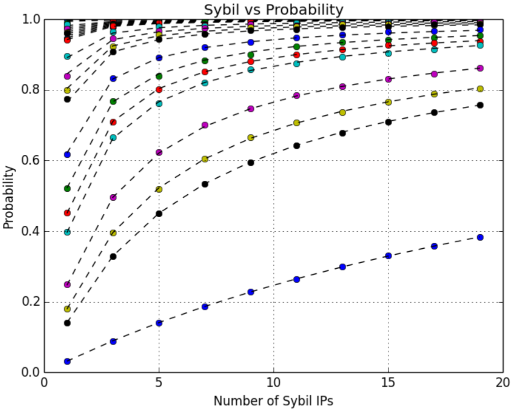

Figure: Our simulation results. The *x*-axis corresponds to the number of IP addresses the adversary can bring to bear. The *y*-axis is the probability that a random region between two adjacent normal members of the network can be mashed. Each line maps to a different network size of *n*. The dotted line traces the line corresponding to the Equation 2: $P_{bad\_neighbor} = \frac{num\_ips \cdot 16383}{num\_ips \cdot 16383 + n - 1}$

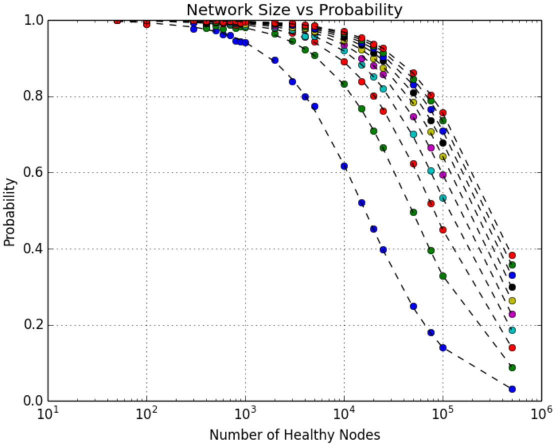

Figure: These are the same as results shown in Figure 1, but our x-axis is the network size n in this case. Here, each line corresponds to a different number of unique IP addresses the adversary has at their disposal.

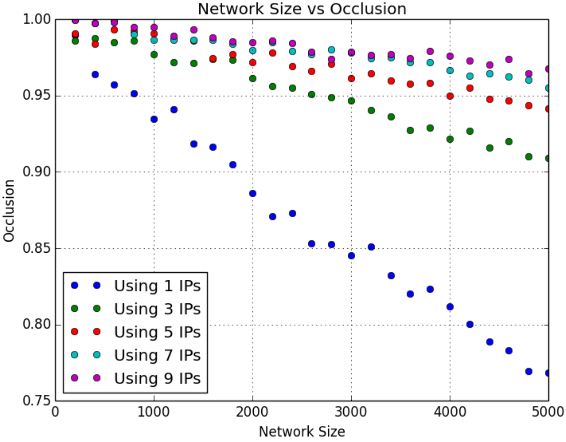

Figure: This graph shows the relationship between the network size and the probability a particular link, adjacent or not, can be mashed.

Questions?

Mihir Bellare and Tadayoshi Kohno.
Hash function balance and its impact on birthday attacks.
In *Advances in Cryptology-Eurocrypt 2004*, pages 401–418. Springer, 2004.

John R Douceur.
The sybil attack.
In *Peer-to-peer Systems*, pages 251–260. Springer, 2002.

Donald Eastlake and Paul Jones.
Us secure hash algorithm 1 (sha1), 2001.