Your Censor is My Censor: Weaponizing Censorship Infrastructure for Availability Attacks

Kevin Bock      Pranav Bharadwaj      Jasraj Singh      Dave Levin
Nation-state censorship
Nation-state censorship

Web browser
Nation-state censorship

Web browser

Website
Nation-state censorship

Web browser

Censor

Website
Nation-state censorship

puppies
Nation-state censorship
Nation-state censorship

kittens
Nation-state censorship
Nation-state censorship

religion
Nation-state censorship
Nation-state censorship

religion
Nation-state censorship

religion
Nation-state censorship
Residual censorship
Block *all* communication for some time
Residual censorship

Block *all* communication for some time

puppies
Residual censorship
Block *all* communication for some time
Residual censorship
Block *all* communication for some time

puppies
Residual censorship

Block *all* communication for some time
Residual censorship
Block *all* communication for some time
Types of residual censorship
Categorized by what information censor remembers

IP:port

IP:port
Types of residual censorship
Categorized by what information censor remembers

4-tuple (IP, port, IP, port)
3-tuple (IP, IP, port)
2-tuple (IP, IP)

More aggressive
Censorship infrastructure
Censorship infrastructure
Censorship infrastructure can be weaponized
Weaponizing residual censorship
Weaponizing residual censorship
Weaponizing residual censorship

religion
Weaponizing residual censorship
Weaponizing residual censorship

religion
Weaponizing residual censorship

religion
Weaponizing residual censorship

religion
Weaponizing residual censorship

religion
Weaponizing residual censorship
Weaponizing residual censorship
Weaponizing residual censorship
Weaponizing residual censorship

puppies
Weaponizing residual censorship

puppies

STOP
Weaponizing residual censorship
Weaponizing residual censorship

Attackers can restrict *benign* communication from crossing the censors’ borders
Weaponizing residual censorship

- How long does it last?
- How can we trigger censors?
- Limitations of residual censorship?
- How can we evaluate ethically?
Weaponizing Residual Censorship

**Current state of residual censorship?**
- Experiments in Iran, China, and Kazakhstan
- Find differences in implementation and duration

**How can it be weaponized?**
- Ethical experiments with SP³
- Attacked ourselves from a dozen vantage points
Weaponizing Residual Censorship

Current state of residual censorship?
- Experiments in Iran, China, and Kazakhstan
- Find differences in implementation and duration

How can it be weaponized?
- Ethical experiments with SP³
- Attacked ourselves from a dozen vantage points
Null routing

<table>
<thead>
<tr>
<th>Diversity of censors</th>
<th>Diversity of protocols</th>
</tr>
</thead>
<tbody>
<tr>
<td>Injects TCP RSTs</td>
<td>HTTP</td>
</tr>
<tr>
<td>China</td>
<td><img src="https://via.placeholder.com/150" alt="China" /></td>
</tr>
<tr>
<td>Iran</td>
<td><img src="https://via.placeholder.com/150" alt="Iran" /></td>
</tr>
<tr>
<td>Kazakhstan</td>
<td><img src="https://via.placeholder.com/150" alt="Kazakhstan" /></td>
</tr>
</tbody>
</table>
## Types of residual censorship

### 3-tuple or 4-tuple

<table>
<thead>
<tr>
<th></th>
<th>HTTP</th>
<th>SNI</th>
<th>ESNI</th>
<th>DNS</th>
<th>SMTP</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>China</strong></td>
<td>3-tuple</td>
<td>3-tuple</td>
<td>both</td>
<td>✗</td>
<td>✗</td>
<td>✗</td>
</tr>
<tr>
<td><strong>Iran</strong></td>
<td>4-tuple</td>
<td>4-tuple</td>
<td>✗</td>
<td>✗</td>
<td>✗</td>
<td>4-tuple</td>
</tr>
<tr>
<td><strong>Kazakhstan</strong></td>
<td>4-tuple</td>
<td>4-tuple</td>
<td>✗</td>
<td>✗</td>
<td>✗</td>
<td>✗</td>
</tr>
</tbody>
</table>
## Duration of residual censorship

How long does blocking last?

<table>
<thead>
<tr>
<th></th>
<th>HTTP</th>
<th>SNI</th>
<th>ESNI</th>
<th>DNS</th>
<th>SMTP</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>China</td>
<td>90s</td>
<td>60s</td>
<td>120s</td>
<td>✗</td>
<td>✗</td>
<td>✗</td>
</tr>
<tr>
<td>Iran</td>
<td>180s</td>
<td>180s*</td>
<td>✗</td>
<td>✗</td>
<td>✗</td>
<td>60s</td>
</tr>
<tr>
<td>Kazakhstan</td>
<td>120s</td>
<td>120s</td>
<td>✗</td>
<td>✗</td>
<td>✗</td>
<td>✗</td>
</tr>
</tbody>
</table>
Is residual censorship bidirectional?
Does it affect traffic entering the country?

<table>
<thead>
<tr>
<th></th>
<th>HTTP</th>
<th>SNI</th>
<th>ESNI</th>
<th>DNS</th>
<th>SMTP</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>China</strong></td>
<td>✓</td>
<td>✓</td>
<td>✓*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Iran</strong></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td>✗</td>
</tr>
<tr>
<td><strong>Kazakhstan</strong></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

In all cases, censor tracks traffic direction.
State of residual censorship

Residual censorship is implemented differently around the world

- Different censorship mechanisms (RSTs vs Null Routing)
- Different types of censorship, even within countries
- Bi-directional, but direction matters
Weaponizing Residual Censorship

**Current state of residual censorship?**
- Experiments in Iran, India, China, and Kazakhstan
- Find differences in implementation and duration

**How can it be weaponized?**
- Ethical experiments with $SP^3$
- Attacked ourselves from a dozen vantage points
Weaponizing Residual Censorship

Current state of residual censorship?
- Experiments in Iran, India, China, and Kazakhstan
- Find differences in implementation and duration

How can it be weaponized?
- Ethical experiments with SP³
- Attacked ourselves from a dozen vantage points
How can we trigger censors?

Issue a request for forbidden content
How can we trigger censors?

Issue a request for forbidden content

Weaponizing Middleboxes for TCP Reflected Amplification

to appear in USENIX Security later this summer
How can we trigger censors?
Issue a request for forbidden content

Weaponizing Middleboxes for TCP Reflected Amplification

Packet sequences

- SYN with Request
- PSH
- PSH+ACK
- SYN ; PSH
- SYN ; PSH+ACK

to appear in USENIX Security later this summer
How can we trigger censors?

Issue a request for forbidden content

Packet sequences:
- SYN with Request
- PSH
- PSH+ACK
- SYN ; PSH
- SYN ; PSH+ACK

Weaponizing Middleboxes for TCP Reflected Amplification to appear in USENIX Security later this summer

Censorship can be triggered without a proper 3-way handshake
How can we trigger censors?

Issue a request for forbidden content

Packet sequences

- SYN with Request
- PSH
- PSH+ACK
- SYN ; PSH
- SYN ; PSH+ACK

Weaponizing Middleboxes for TCP Reflected Amplification
to appear in USENIX Security later this summer

Censorship can be triggered without a proper 3-way handshake
Ethical evaluation

Attack ourselves *ethically* without affecting other hosts

- Need to spoof traffic to or from a censored regime
- Only between hosts we control
- Full control over packets we send

Solution: SP³
Experiment Setup with SP³
Experiment Setup with SP³

- **Client**
- **Censor**
- **Server**
- **SP³**

Consent to receive source-spoofed packets
Experiment Setup with SP³

- Client
- Server
- Censor
- SP³
Experiment Setup with SP$^3$
Experiment Setup with SP³

Censor
Experiment Setup with SP³

Censor

Innocuous request
Experiment Setup with SP³

Around the world

Innocuous request

Censored Regimes

University of Washington

China

Kazakhstan

Iran
Experiment Setup with SP³

“Attacker” (SP³)

“Victims” (Our clients)

Censoring Nation-states

“Victims” (Our servers)

Tested from 16 external vantage points
Experiment Setup with SP³

Tested from 16 external vantage points

“Attacker” (SP³) — “Victims” (Our clients) — “Victims” (Our servers)

Censoring Nation-states
### Results

<table>
<thead>
<tr>
<th>Victim Location</th>
<th>Destination Location</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Kazakhstan</td>
</tr>
<tr>
<td>Australia</td>
<td>Sydney</td>
</tr>
<tr>
<td>China</td>
<td>Beijing 1</td>
</tr>
<tr>
<td></td>
<td>Beijing 2</td>
</tr>
<tr>
<td>India</td>
<td>Mumbai</td>
</tr>
<tr>
<td></td>
<td>Bangalore 1</td>
</tr>
<tr>
<td></td>
<td>Bangalore 2</td>
</tr>
<tr>
<td>Iran</td>
<td>Tehran</td>
</tr>
<tr>
<td>Ireland</td>
<td>Dublin 1</td>
</tr>
<tr>
<td></td>
<td>Dublin 2</td>
</tr>
<tr>
<td>Japan</td>
<td>Tokyo</td>
</tr>
<tr>
<td>Kazakhstan</td>
<td>Qaraghandy</td>
</tr>
<tr>
<td>Russia</td>
<td>Khabarovsk</td>
</tr>
<tr>
<td>UAE</td>
<td>Dubai 1</td>
</tr>
<tr>
<td></td>
<td>Dubai 2</td>
</tr>
<tr>
<td>USA</td>
<td>Colorado</td>
</tr>
<tr>
<td></td>
<td>Iowa</td>
</tr>
<tr>
<td></td>
<td>Virginia</td>
</tr>
<tr>
<td>Victim Location</td>
<td>Destination Location</td>
</tr>
<tr>
<td>-----------------</td>
<td>----------------------</td>
</tr>
<tr>
<td></td>
<td>Kazakhstan</td>
</tr>
<tr>
<td>Australia</td>
<td>Sydney</td>
</tr>
<tr>
<td>China</td>
<td>Beijing 1</td>
</tr>
<tr>
<td></td>
<td>Beijing 2</td>
</tr>
<tr>
<td>India</td>
<td>Mumbai</td>
</tr>
<tr>
<td></td>
<td>Bangalore 1</td>
</tr>
<tr>
<td></td>
<td>Bangalore 2</td>
</tr>
<tr>
<td>Iran</td>
<td>Tehran</td>
</tr>
<tr>
<td>Ireland</td>
<td>Dublin 1</td>
</tr>
<tr>
<td></td>
<td>Dublin 2</td>
</tr>
<tr>
<td>Japan</td>
<td>Tokyo</td>
</tr>
<tr>
<td>Kazakhstan</td>
<td>Qaraghandy</td>
</tr>
<tr>
<td>Russia</td>
<td>Khabarovsk</td>
</tr>
<tr>
<td>UAE</td>
<td>Dubai 1</td>
</tr>
<tr>
<td></td>
<td>Dubai 2</td>
</tr>
<tr>
<td>USA</td>
<td>Colorado</td>
</tr>
<tr>
<td></td>
<td>Iowa</td>
</tr>
<tr>
<td></td>
<td>Virginia</td>
</tr>
</tbody>
</table>

- **Results**: The table shows the HTTP and HTTPS status for different locations and destinations, along with the percentage of ESNI usage. For example, in Sydney, all connections are via HTTPS, and in Iran, Tehran has 50% ESNI usage.
Why does it fail?

Source-spoofed traceroute from both to compare network path
Why does it fail?

Source-spoofed traceroute from both to compare network path
Why does it fail?

Source-spoofed traceroute from both to compare network path
Why does it fail?
Why does it fail?
Why does it fail?

Different entry points

STOP
Sustaining the attack
Goal: block client IP to server IP:port

Depends on type of residual censorship
Sustaining the attack

Goal: block client IP to server IP:port

Client

Censor

Server

Attacker can’t guess source port

4-tuple (IP, port, IP, port)

3-tuple (IP, IP, port)
Sustaining the attack: 4-tuple

Goal: block client IP to server IP:port

Attacker can re-trigger from all 65,535 src ports
Sustaining the attack: 4-tuple

Goal: block client IP to server IP:port

Attacker can re-trigger from all 65,535 src ports
Sustaining the attack: 4-tuple

Goal: block client IP to server IP:port

Speed required = \frac{\text{Trigger packets}}{\text{Duration}} \times 65,535
Sustaining the attack: 4-tuple

Goal: block client IP to server IP:port

\[
\text{Speed required} = \frac{\text{Trigger packets}}{\text{Duration}} \times 65,535 = \frac{145 \text{ bytes}}{120 \text{ seconds}} \times 65,535 = 634 \text{ kbps}
\]
Sustaining the attack: 3-tuple
Goal: block client IP to server IP:port

Weak attacker can launch this attack effectively
Sustaining the Attack
Victim helps sustain the attack
Sustaining the Attack
Victim helps sustain the attack

Censor

Innocuous request
Sustaining the Attack
Victim helps sustain the attack

Residual timer resets if the victim sends data

Victim retransmissions unknowingly sustain the attack on themselves
Can the server detect this?

Attacker can limit TTL of packets to reach censor, but not server
Can the server detect this?

Attacker can limit TTL of packets to reach censor, but not server
Attack Limitations

Attacker must have a vantage point:

1. Without egress filtering
2. Shares a similar enough path with their victim
3. Traffic crosses a censor (with residual censorship)
4. Censor can be triggered statelessly

Surprisingly high number of shared network paths
What can be done?

Some mitigations available to censorship infrastructure:

- Abolish 3-tuple residual censorship
- Null routing should track sequence numbers
- Properly track 3-way handshake

Unfortunately, no good countermeasures available to victims
Other details in the paper

- **Port Experiments**: Examine which ports are affected
- **Reliability Experiments**: Studied the reliability of residual censorship
- **ESNI Weaponization**: Details on how to weaponize ESNI in China
- **Attack Breadth**: Analysis of other countries that might be affected
Weaponizing Middleboxes

Censors can be weaponized to launch availability attacks

Can be done from a weak attacker

Censors pose a threat to the entire Internet

Code and website censorship.ai