Ignoring the Great Firewall of China

Richard Clayton,
Steven J. Murdoch, Robert N.M. Watson

UNIVERSITY OF CAMBRIDGE
Computer Laboratory

Berkeley, CA

26th July 2006

Summary

- Content blocking system taxonomy
- The “Great Firewall of China”
- Ignoring the Chinese firewall
- Denial of Service attacks
- Chinese firewall design
- Firewall SYN/ACK confusion
- Conclusions

Content blocking methods

- Blackhole routeing of IP addresses
  - fine for major sites, but collateral damage possible & have to keep database updated
- DNS poisoning (do not provide IP address)
  - fine for major sites, updating also a problem
- Use web proxy to filter if URL match
  - expensive at country scale, at a time when web proxy caches are going out of fashion

Keyword filtering

- Chinese firewall shuts connections if it spots specific keywords passing by
  - for example **`GET /?falun HTTP/1.0`**
- Keywords spotted as they pass by in both directions (dealing with requests & results)
- ***CAUTION:*** parts of Chinese system DO use other blocking methods, and the academic network isn't currently using the scheme, and other protocols are blocked at the application level!

Actual mechanism

```
cam(54190) → china(http)[SYN]
china(http)→ cam(54190) [SYN, ACK] TTL=39
cam(54190) → china(http)[ACK]
cam(54190) → china(http) GET /?falun HTTP/1.0<crlf><crlf>
china(http)→ cam(54190) [RST] TTL=47, seq=1, ack=1
china(http)→ cam(54190) [RST] TTL=47, seq=1461, ack=1
china(http)→ cam(54190) [RST] TTL=47, seq=4381, ack=1
china(http)→ cam(54190) HTTP/1.1 200 OK (text/html)<crlf>..
cam(54190) → china(http)[RST] TTL=64, seq=25, ack zeroed
china(http)→ cam(54190) . . . more of the web page
cam(54190) → china(http)[RST] TTL=64, seq=25, ack zeroed
china(http)→ cam(54190) [RST] TTL=47, seq=2921, ack=25
```

Meanwhile…

- The other end of the connection is ***also*** seeing RST packets from the firewall!

Ignoring the firewall

- **Q:** Since the packets pass through the firewall, what happens if the RST packets are ignored?
- **A:** Web page is transferred just fine (though you get a LOT more RSTs as well)
- NB: necessary to ignore RST packets at ***both*** ends of the connection

Further connections

- Trying to connect again causes RST packets to be sent immediately (even if no “bad” keywords are transferred)

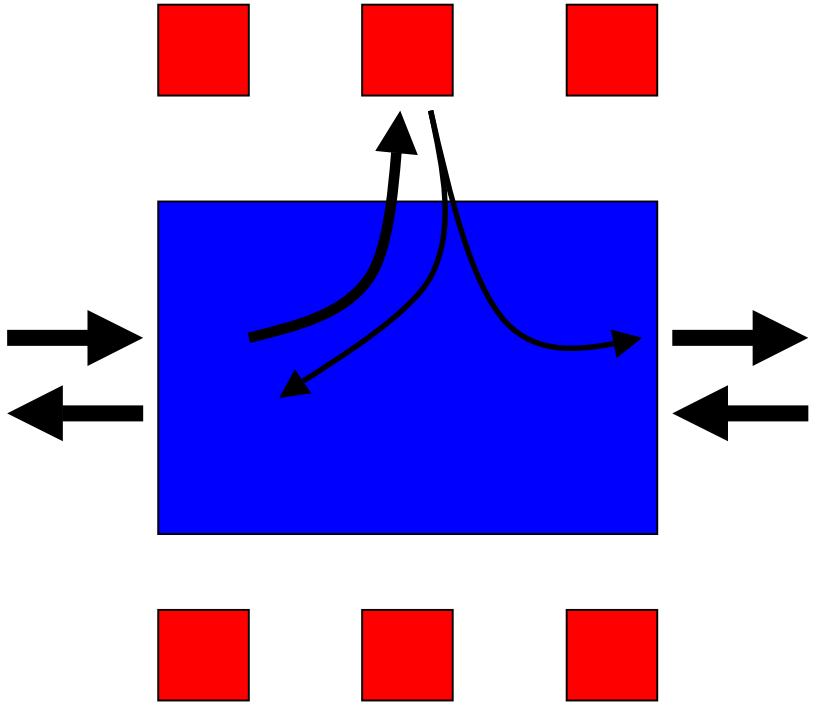
```
cam(54191) → china(http)[SYN]
china(http)→ cam(54191) [SYN, ACK] TTL=41
cam(54191) → china(http)[ACK]
china(http)→ cam(54191) [RST] TTL=49, seq=1
```

- Once again dropping RSTs allows transfer

Denial of service attack

- Send single packets (containing `falun`) to Chinese firewall, forging source & destination
- Connection from source to destination blocked
- Single dialup connection can knock many hundreds of connection over
- NB: only pairs of addresses
- NB: only nearby port numbers (? NAT ?)

Firewall design



Evidence:

- RST sometimes precedes & sometimes follows data
- RST values (+0, +n, +3n)
- Read the user manuals from (?)providers
- Shuffling of RSTs when a sudden burst of packets

NB: NO STATE IN FIREWALL!

Firewall “state”?

- Splitting `?falun` across packets avoids detection (a surprise! hardware thought to be used can handle this (and overlaps!))
- Refined view is that firewall doesn’t assume it sees packets in both directions, so must do the best it can with the packet in its hand
- Future work will refine our explanation

False SYN/ACKs

```
cam(38104) → china(http)[SYN]
china(http)→ cam(38104) [SYN, ACK] TTL=105
cam(38104) → china(http)[ACK]
cam(38104) → china(http) GET / HTTP/1.0<crlf><crlf>
china(http)→ cam(38104) [RST] TTL=45, seq=1
china(http)→ cam(38104) [RST] TTL=45, seq=1
china(http)→ cam(38104) [SYN, ACK] TTL=37
cam(38104) → china(http)[RST] TTL=64, seq=1
china(http)→ cam(38104) [RST] TTL=49, seq=1
china(http)→ cam(38104) [RST] TTL=45, seq=3770952438
china(http)→ cam(38104) [RST] TTL=45, seq=1
china(http)→ cam(38104) [RST] TTL=45, seq=1
china(http)→ cam(38104) [RST] TTL=45, seq=1
china(http)→ cam(38104) [RST] TTL=45, seq=1
```

Fixing “blocking with confusion”

- Fake SYN/ACK does not confuse once real SYN/ACK has been accepted
- SYN/ACK ***currently*** easy to distinguish
- Real fix is for stack to hold alternative views of remote sequence value, avoid using a value until see further evidence
  - lack of state in firewall makes this easy(ish)

Porn vs Politics

- Firewall capable of logging events
- No different from encryption/proxies – **but** firewall knows if you're looking at porn or at politics: so may affect your sentence
- Special code is evidence on your machine
- Much better if stack vendors made special tools unnecessary; and there's technical reasons to wish to drop fake resets

Conclusions

- A key part of the Great Firewall of China relies on acquiescence by the end-points
  - more MitM (such as SYN/ACK) possible
- Evasion requires (in)action at both ends
- Firewall can still log exceptions
  - but can distinguish porn from politics
- Stack vendors could provide standard fix
- Other systems may be vulnerable (& to DDoS)

Thanks

Assistance was provided for logging etc by a Chinese citizen [who was unaware of what we proposed to do]. Their site does NOT contain any material that should be censored and no censorable requests were made from the Chinese end of the connection.

Ignoring the Great Firewall of China

Richard Clayton,
Steven J. Murdoch, Robert N.M. Watson

`http://www.cl.cam.ac.uk/~rnc1/ignoring.pdf`

`http://www.lightbluetouchpaper.org/`

UNIVERSITY OF CAMBRIDGE
Computer Laboratory