# Statement of Avi Rubin
# China's Cyber-Wall: Can technology break through?
# November 4, 2002

While I am a researcher at AT&T Labs, I am participating in this round table as an individual, representing only my personal beliefs and opinions. I have been researching computer security issues since 1991. Much of my work has focused on privacy, anonymity, and censorship resistance.

The purpose of my statement is to discuss technical issues related to censorship. I will discuss the techniques that a network administrator, including a large company or a country, could use to censor access and content to and from its network, and I will discuss techniques that could be used to circumvent this censorship. For the remainder of this paper, I will refer to the party controlling the network as the Censor, and to the party wishing to circumvent censorship as the User.

Censorship is somewhat of a broad term. It can refer to the *blocking* of access to web sites. It can refer to *blocking* all connectivity outside of the domain of the Censor, and censorship can refer to the *limitation of access* to certain content. Censorship can also involve *forceful removal* of content from the Web, by applying pressure to the publisher and/or the web hosting party. The latter is the type of censorship that the Publius system was designed to circumvent. In this statement, I do not discuss censorship within the domain of the Censor, but rather, the censorship of content available from outside of the domain for people whose network is under the control of the Censor. I also focus on the User as the receiving party of information and not the publishing party. I will be happy to discuss issues related to the latter in the question and answer period.

There are three principle techniques that can be employed by the Censor.

1. **Routing filters:** The Censor is in a position to control how traffic from the User reaches the rest of the Internet. The Censor can refuse to route Internet packets from the User that are destined for particular locations. Thus, the Censor can use the destination address of the packets to make a censorship decision. In the extreme, the Censor can prevent all traffic from all of its users from reaching any network outside of its control. This is easy to do, and any Censor can accomplish this without the need to purchase any new hardware or software. The functionality is built into all off the shelf routing equipment that sites use to connect to the Internet.

2. **DNS tricks:** The Censor can exert some control on which external sites users can communicate with by virtue of its control over the Domain Name Servers (DNS) within its administrative boundary. The DNS is the service that maps computer addresses (IP addresses) to names. For example, www.avirubin.com has the address 207.140.168.155. Computers communicate using such numerical address, but people enter readable names into web browsers. The DNS translates these names into numbers. Since the Censor controls its own DNS service, it can translate requests from the User to addresses under its own control. For example, if the User attempts to connect to www.avirubin.com, the Censor can program its DNS to return 10.10.32.1 when the User's machine tries to figure out the IP address of the machine, and this address can be that of a machine controlled by the Censor. Thus, DNS provides the Censor with the ability to control which computers the User can connect to.

3. **Application level filtering:** The previous censorship techniques dealt specifically with connectivity issues. Application level filtering, on the other hand, is a mechanism for controlling the content, even if the User can connect to a server. The most likely type of application level filter that the Censor would use is an HTTP proxy. This is a program that intercepts requests sent to Web servers and the responses returned to the User. The Censor can inspect the content, and a decision can be made, as to whether or not to block the information from reaching the User. A Censor using an HTTP proxy might focus its attention on popular search engines.

The first type of censorship, based on routing filters, is difficult to circumvent. If the routers do not allow packets in and out of the network, then there is no way to get around that. The best one could do is to dial up to an external ISP. Of course, this could get expensive if the Censor is a country. Also, a very strict and powerful censor could monitor the phone network for data dial-up connections and disconnect them, as well as sanction the User.

The second type of censorship, based on DNS spoofing, can be circumvented by users who know the IP address of the server with which they wish to communicate. Instead of referring to the server by name, they could connect using the IP address directly. However, IP addresses change frequently, and it may not always be possible for users under the control of the Censor to know the IP address of a server. In general, this is not a very effective technique.

The third type of censorship, based on application level filtering, is perhaps the easiest to circumvent. Encrypted content is difficult to censor, but a very strict Censor can maintain a policy of blocking all content that it cannot interpret for the purposes of filtering. Perhaps the easiest way to bypass HTTP proxies is to proxy web content over a different port. Port numbers are used on the Internet to identify the type of service for packets between hosts. For example, Web traffic uses port 80. HTTP proxies process packets that are marked with port 80. A User wishing to circumvent this monitoring could cooperate with someone on the outside of the Censor's administrative control. They could set up two proxies. The inside one would translate port 80 packets into ones that use, say, port 14500. The outside one would translate port 14500 back to port 80 and send them to the server. Thus, the User could browse the Web without the Censor detecting it. However, a strict censor could block all ports except 80, and then filter on port 80. There is little that could be done by the User in that case. It should be noted that researchers have succeeded in identifying services by their traffic patterns, independent of port numbers.

The bottom line is that there is an arms race in censorship. An extreme Censor can win every time, but at the expense of completely disconnecting all users. The more tolerant a Censor, the more avenues there will be for circumvention of the censorship that is in place.