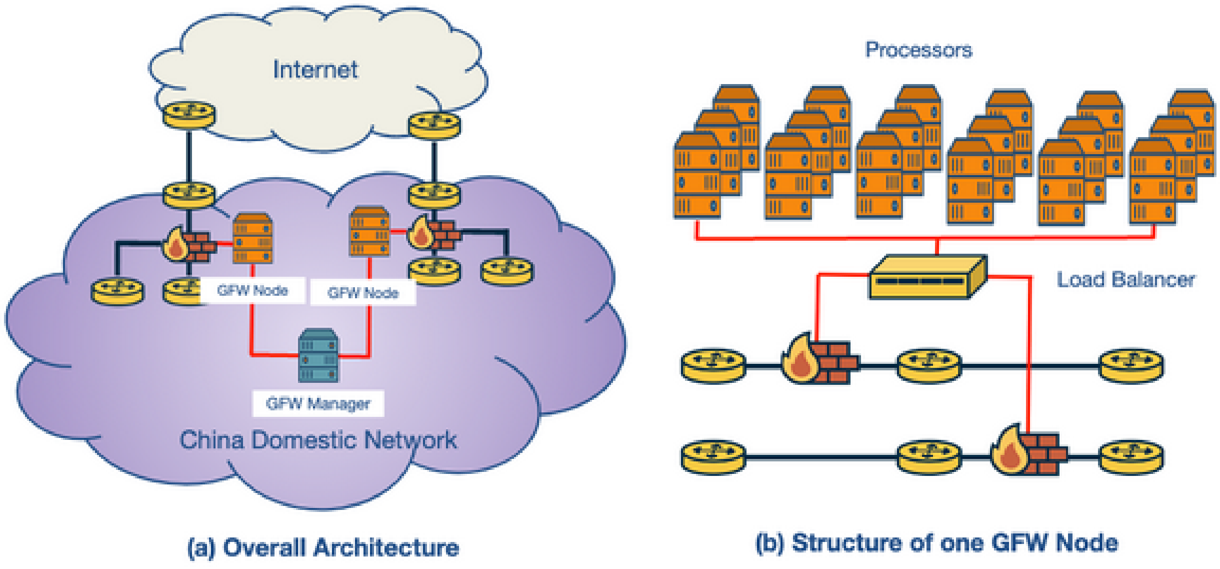
Internet Surveillance and Censorship

Introduction

1. Internet Censhorship is a special case of Internet security
   - More subtle category of attacks and presents its own unique challenges to detect and measure it
   - Cover techniques attackers have developed to abuse popular protocols such as DNS and BGP to gain access to information
   - Three types of censorship:
     - Internet connectivity
     - DNS censorship
     - Social media-based censorship

DNS Censorship: What is it?

1. What is DNS Censorship?
   - Large scale network traffic filtering strategy opted by a network to enforce control and censhorship over Internet infrastructure to suppress material which they deem as objectionable
     - Great Firewall of China is a firewall that uses various techniques to censor China's Internet traffic and block access to foreign websites
2. Great Firewall of China
   - Works by injecting fake DNS record responses so that access to a domain name is blocked
   - Several studies conducted to deduce the functionality of the system
   - Properties of GFW
     - Locality of GFW Nodes: Unclea where GFW nodes are present only at the edge ISPs or whether they are also present in non-bordering Chinese ASes
       - Majoring view is present at the edge
     - Centralized management: Blocklists obtained from two distinct GFW locations are the same, so there is a high probability of a central management (GFW Manager) entity that orchestrates blocklists
     - Load balancing: GFW load balances between processes based on source and destination IP address. Processes are clustered together to collectively send injected DNS responses



(a) Overall Architecture

(b) Structure of one GFW Node

Structure of One GFW Node

Example DNS Censorship Techniques (1)

1. How does DNS Injection work?
    - One of the most common censorship techniques employed by the GFW
    - GFW uses a ruleset to determine when to inject DNS replies to censor network traffic
        - Important to identify and isolate the networks that use DNS Injection for censorship
    - Accuracy of DNS open resolvers to accurately pollute the response is recorded over 99.9%
    - Steps involved in DNS injection:
        - DNS probe is sent to the DNS resolvers
        - Probe is checked against the blocklist of domains and keywords
        - For domain level blocking, a fake DNS A record response is sent back.
            - Can block domain directly
            - Can also block based on keywords present in the domain

Example DNS Censorship Techniques (2)

1. Packet Dropping
    - All network traffic going to a set of specific IP addresses is discarded
    - Censor identifies undesirable traffic and chooses not to properly forward any packets it sees associated with the traversing undesirable traffic instead of following a normal routing protocol
    - Strengths
        - Easy to implement
        - Low cost
    - Weaknesses
        - Maintenance of blocklist: Challenging to stay up to date
        - Overblocking: If two websites share an IP address and the intent is to block one, there's a risk of blocking both
2. DNS Poisoning
    - When a DNS receives a query for resolving hostname to IP address but the server returns no answer or an incorrect answer to redirect or mislead the user request
    - Strengths
        - No overblocking: Since there is an extra layer of hostname translation, access to specific hostnames can be blocked versus blanket IP address blocking
        - Blocks the entire domain; not possible to allow email contact while blocking the website
3. Content Inspection
    - Proxy-based content inspection: All network traffic passes through to a proxy where the traffic is examined for content and rejects requests that serve objectionable content
    - Strengths
        - Precise censorship: Can censor down to single web pages or objects within a web page
        - Flexible: Works well with hybrid security systems
    - Weaknesses
        - Not scalable: Expensive to implement on a large scale network as the processing overhead is large
    - IDS-based content inspection: Alternative approach is to use parts of an IDS to inspect network traffic. IDS is easier and more cost effective to implement than a proxy-based system as it is more responsive than reactive in nature by informing firewall rules for censorship
4. Blocking with Resets
    - Send a TCP reset (RST) to block individual connections that contain requests with objectionable content
5. Immediate Reset of Connections
    - Censorship systems like GFW have blocking rules in addition to inspecting content, to suspend traffic coming from a source immediately, for a short period of time

Quiz 1

1. The Great Firewall of China injects fake DNS A records to block individual connections.
   - False
2. The Great Firewall of China is likely managed by a single entity.
   - True
3. The Great Firewall of China may block content based on which of the following characteristics?
   - Keywords within the URL
   - Images on a webpage
   - Destination IP
4. Packet dropping is a scheme used to censor content. Which of the following statements characterize packet dropping? Select all that apply.
   - Low cost to implement
   - Might block content otherwise deemed appropriate
5. The GFW can block a portion of a website using DNS poisoning.
   - False
6. Suppose a client in Cambridge makes a request to a website based in China. When does the GFW reset the connection?
   - After the ACK sent by the client in Cambridge

Why is DNS Manipulation Difficult to Measure?

1. Challenges to understanding censorship
   - Diverse measurements: Need a diverse set of measurements spanning different geographic regions, ISPs, countries, and regions within a single country
     - Different organizations may implement censorship at multiple layers of the Internet protocol stack and using different techniques
     - Need widespread longitudinal measurements to understand global Internet manipulation and the heterogeneity of DNS manipulation across countries, resolvers, and domains
   - Need for scale: Need for methods and tools that are independent of human intervention and participation
     - Early methods relied on volunteers who were running measurement software on their own devices
   - Identifying the intent to restrict content access: DNS manipulation requires that we detect the intent to block access to content, not just DNS misconfiguration
   - Ethics and minimizing risks: Risks associated with involving citizens in censorship measurement studies based on potential penalties
     - Safer to rely on open DNS resolvers that are hosted in Internet infrastructure, for example, within Internet service providers or cloud hosting providers

Example Censorship Detection Systems and their Limitations

1. Censorship Measurement Tools
   - Created by efforts to measure censorship by running experiments from diverse vantage points
   - OpenNet Initiative had volunteers perform measurements on their home networks at different times since the past decade
     - Made continuous and diverse measurements very difficult
   - Augur is a new system created to perform longitudinal global measurements using TCP/IP side channels

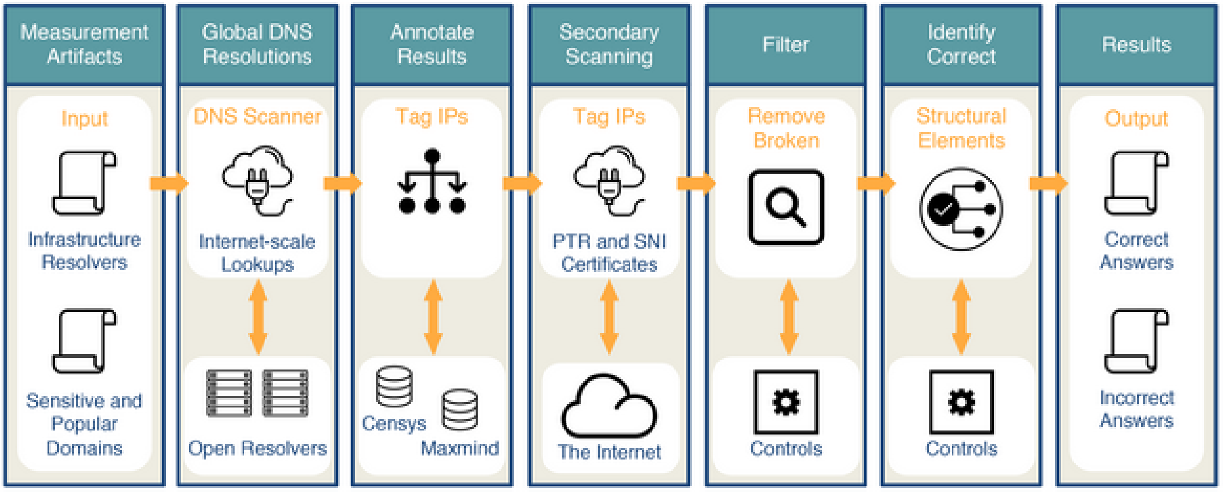
Quiz 2

1. Consider the variance of censorship methods used. Select the statement which correctly describes the situation.
   - Censorship methods are inconsistent across ISPs making it difficult to measure DNS manipulation.

2. Current research methods for understanding DNS methods are scalable due to the number of volunteers participating.
    - False
3. The use of Open DNS resolvers resolves some of the ethical concerns associated with Internet censorship studies.
    - True

DNS Censorship: A Global Measurement Methodology

1. Iris
    - Uses open DNS resolvers located all over the globe to avoid using home routers
    - Two main steps:
        - Scanning the Internet's IPv4 space for open DNS resolvers
        - Identfiying infrastructure DNS resolvers
    - Now that we've obtained a global set of open DNS resolvers, need to perform the measurements
        - Performing global DNS queries: Iris queries thousands of domains across thousands of open DNS resolvers
        - Annotating DNS responses with auxiliary information: To enable the classification, Iris annotates the IP addresses with additional information such as their geo-location, AS< port 80 HTTP responses, etc.
        - Additional PTR and TLS scanning: One IP address could host several websites via virtual hosting, so Iris adds PTR and SNI certificates
2. Consistency Metrics
    - Domain access should have some consistency, in terms of network properties, infrastructure, or content, even when accessed from different global vantage points
    - Metrics: IP address, AS, HTTP content, HTTPS certificate, PTRs for CDN
3. Independent Variability Metrics
    - Also use metrics that could be externally verified using external data sources.
    - Metrics: HTTPS certificate and HTTPS certificate with SNI
    - If any consistency or independent verifiability metric is satisfied, the response is correct; otherwise, it's classified as manipulated



Overview of DNS Resolution, Filtering, and Classification

Quiz 3

1. Iris uses Open DNS resolvers obtain a dataset for machine learning.
2. Suppose Iris is being used to detect DNS manipulation. Iris queries a global resolver for an IP addresses (consistency metric) and receives a DNS A record with a different IP address than the ones stored. Which of the follow statements are true?
    - The response is inconsistent, but might not be classified as manipulated.

Censorship Through Connectivity Disruptions

1. Connectivity Disruptions
    - Highest level of Internet censorship is to completely block access to the Internet
    - More subtle approach is to use software to interrupt the routing or packet forwarding mechanisms
        - Routing disruption: Routers use BGP to communicate updates to other routers in the network and decide which parts of the network are reachable. If this communication is disrupted or disabled on critical routers, it could result in unreachability of the large parts of a network.
            - Easily detectable
        - Packet filtering: Used as a security mechanism in firewalls and switches. Can also be used to block packets matching a certain criteria disrupting the normal forwarding action.
            - Harder to detect and might require active probing of the forwarding path or monitoring traffic of the impacted network

Connectivity Disruptions: A Case Study

1. Egypt
    - On 25th of January, access to Twitter was blocked in response to political developments
    - Complete Internet shutdown on the 27th of January
    - All routes to Egyptian networks were withdrawn from the Internet's global routing table
2. Libya
    - On February 17th, YouTube was blocked
    - On March 3rd, Internet access was disabled completely for 4 days
    - A single AS owned by the state dominates Libya's Internet infrastructure, with only two submarine cables providing international connectivity
    - 12 of the 13 delegated prefixes to Libya were withdrawn by its local telecom operator with a reasonable exception of a prefix that was controlled by any outsidec company

Connectivity Disruptions: Detection

1. Augur
    - Uses a measurement machine to detect filtering between hosts
    - Aims to detect if filtering exists between two hosts, a reflector and a site
    - Reflector is a host which maintains a global IP ID
    - Site is a host that may be potentially blocked
2. IP ID
    - Strategy used by Augur takes advantage of the fact that any packet that is sent by a host is assigned a unique 16-bit IP identifier, which the destination host can use to reassemble a fragmented packet
    - IP ID is different for packets generated by the same host
    - Typically performed by maintaining a global counter
    - Can determine how many packets are generated by a host
3. Probing
    - Mechanism to monitor the IP ID of a host over time
    - Use the measurement machine to observe the IP ID generated by the reflector
4. Perturbation
    - Mechanism which forces a host to increment its IP ID counter by sending traffic from different sources such that the host generates a response packet

- Measurement machine sends a spoofed TCP SYN packet to the site with source address set to the reflector's IP address
- Site responds to the reflector with a TCP SYN-ACK packet
- Reflectro returns a TCP RST packet to the site while also incrementing its global IP ID counter by 1

Quiz 4

1. Augur is used to identify DNS-based manipulations.
   - False
2. Suppose we are using Augur to detect filtering between two host, and that we have a scenario where no blocking occurs. The measurement machine sends a SYN-ACK to the reflector. What should happen?
   - The return IP ID from the reflector to the measurement machine should increase by 2.