# A Deep Dive into DNS Query Failures

**Donghui Yang**[1,2], Zhenyu Li[1], Gareth Tyson[3]

[1]Institute of Computing Technology, Chinese Academy of Sciences

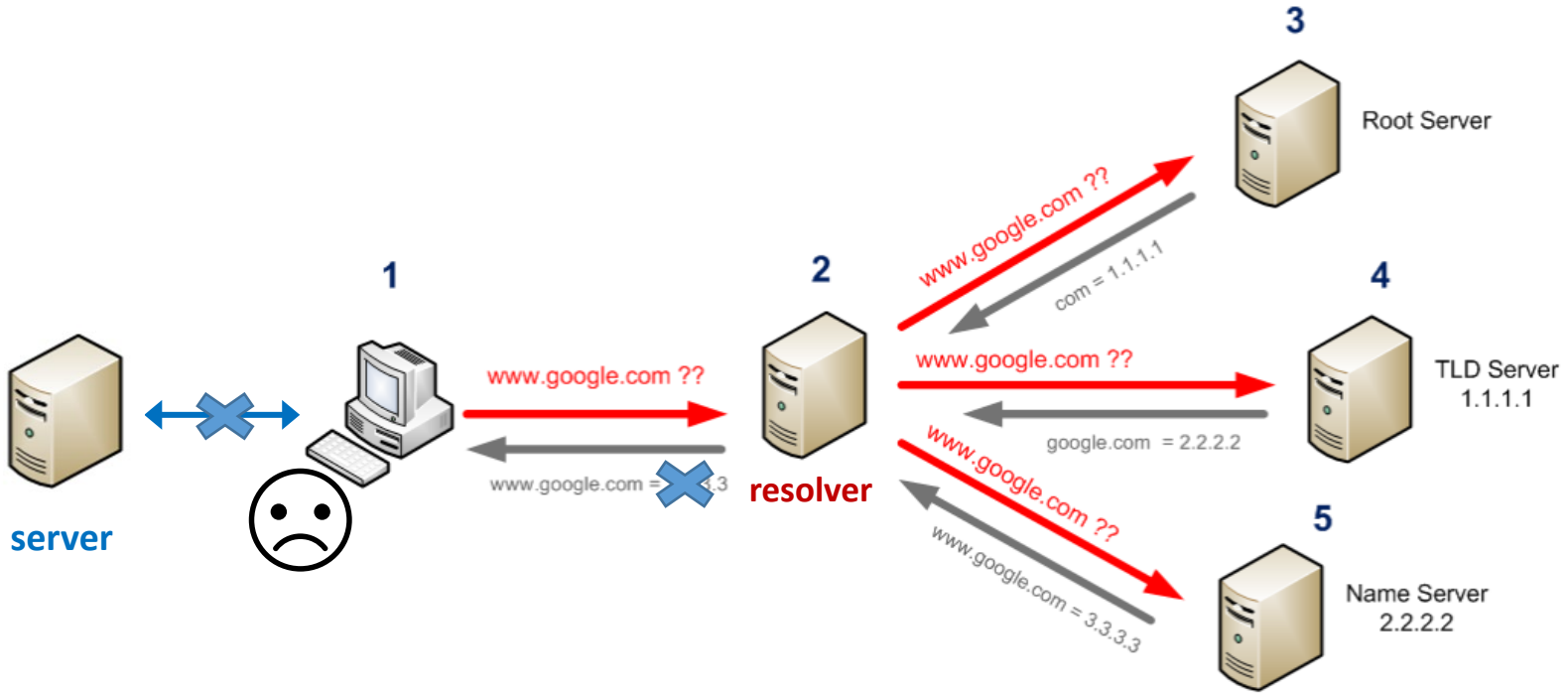
[2]University of Chinese Academy of Sciences

[3]Queen Mary University of London

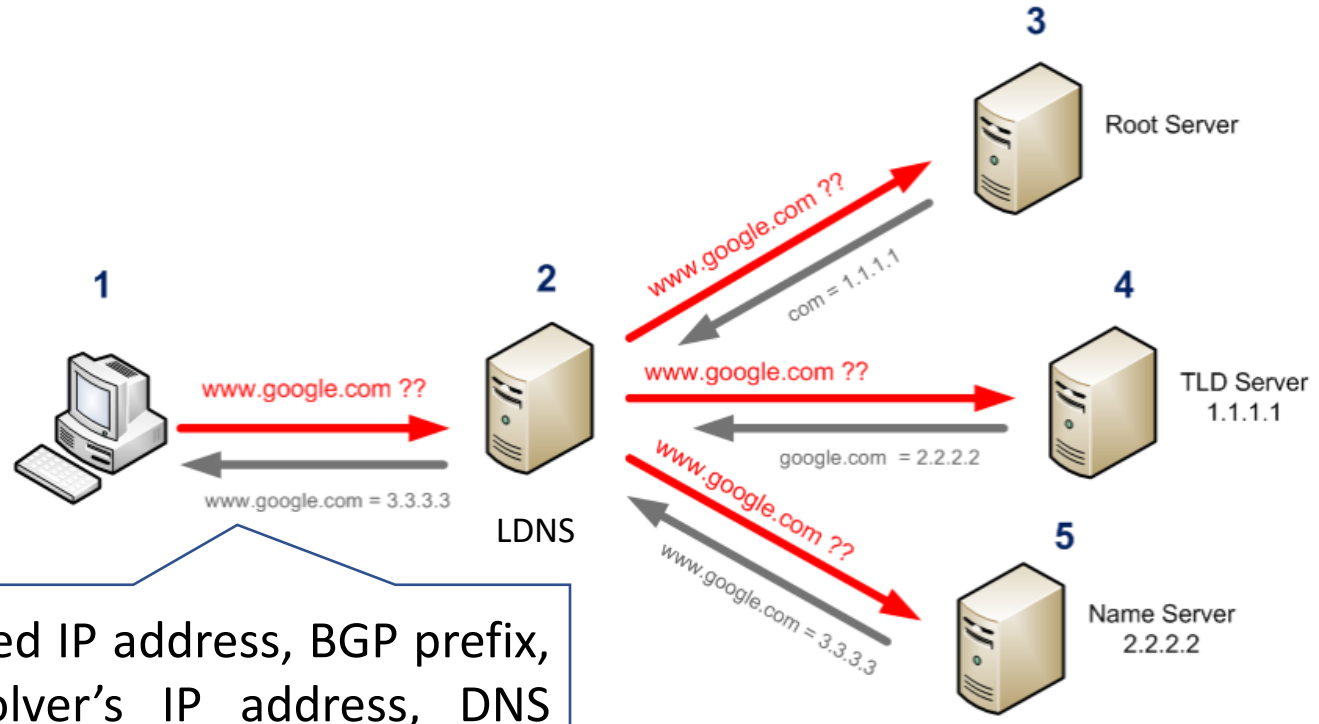


# Why to study DNS Query Failures

- Failures prevent access to any services dependent on domain names
- High-level observation: **13.5%** of DNS queries fail



# Passive DNS Data




end user's anonymized IP address, BGP prefix, ASN, recursive resolver's IP address, DNS query type, resource records, timestamp

- 14-day samples (each sample consists of 10-minute logs) , ~3.1 billion logs

# Identification of Failed Queries

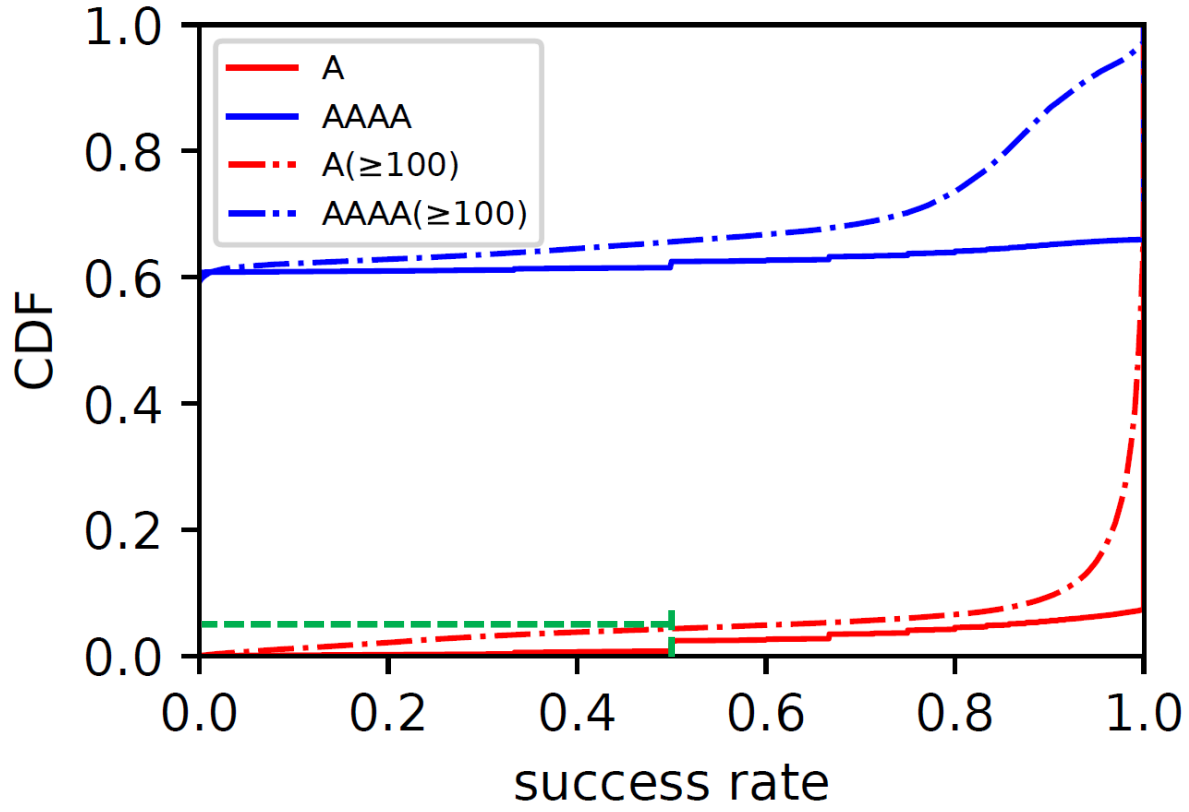
- No RCODE: we turn to a heuristic method to filter out logs that are attributed to NXDOMAINs
- Check if the requested domain (QNAME) contains a *valid* answer
  - e.g., for an A query, at least one RR in the response is an A record of the QNAME
- Extract failed queries of the four most popular types of records that constitute 99.5% of all queries
- Filter out logs attributed to NXDOMAINs by removing logs containing domains that have never succeeded in the whole dataset
  - 2.8 billion logs remain for subsequent analyses

# A Primer on DNS Failures

| Query Type | A | AAAA | PTR | MX | Others |
|---|---|---|---|---|---|
| #queries | 86.2% | 10.4% | 2.8% | 0.1% | 0.5% |
| Success Rate | 93.1% | 35.8% | 40.4% | 82.9% | - |

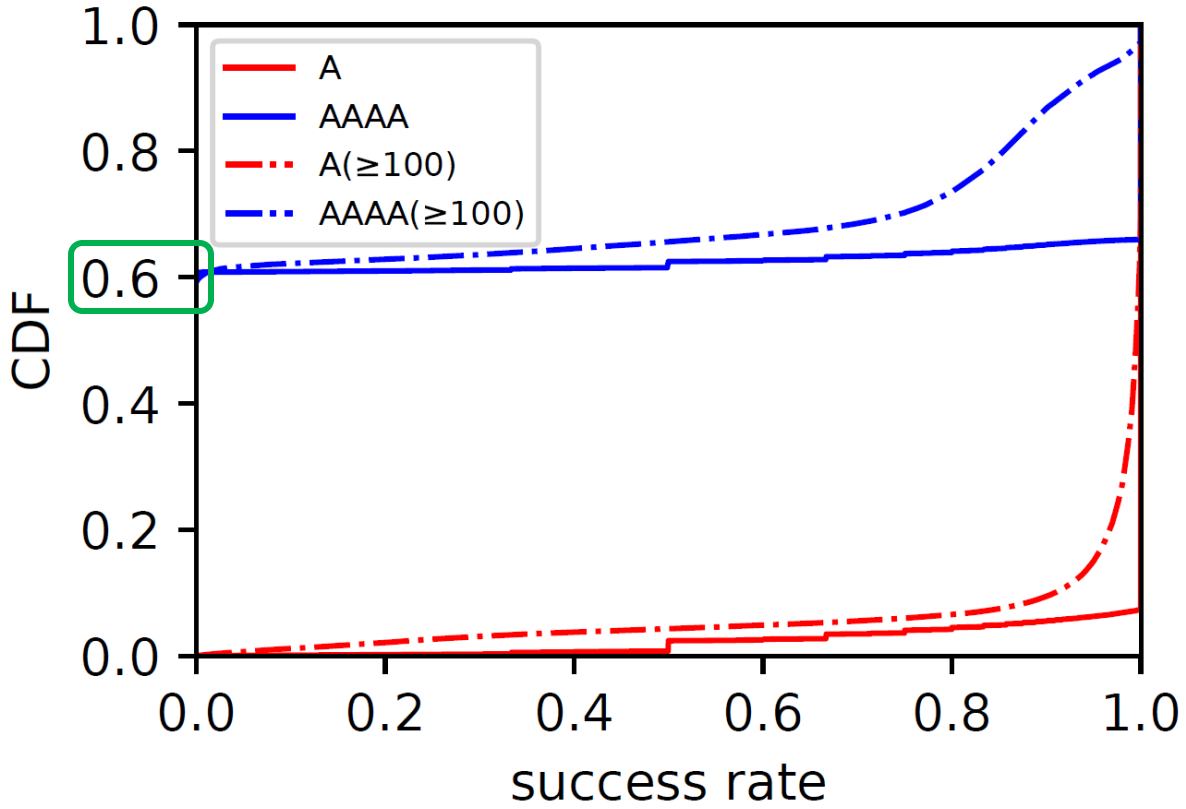
- A queries account for the majority and are successfully resolved most frequently
- Other query types manifest lower success rates
  - Surprisingly low success rate for AAAA queries

# Failures Across Domains



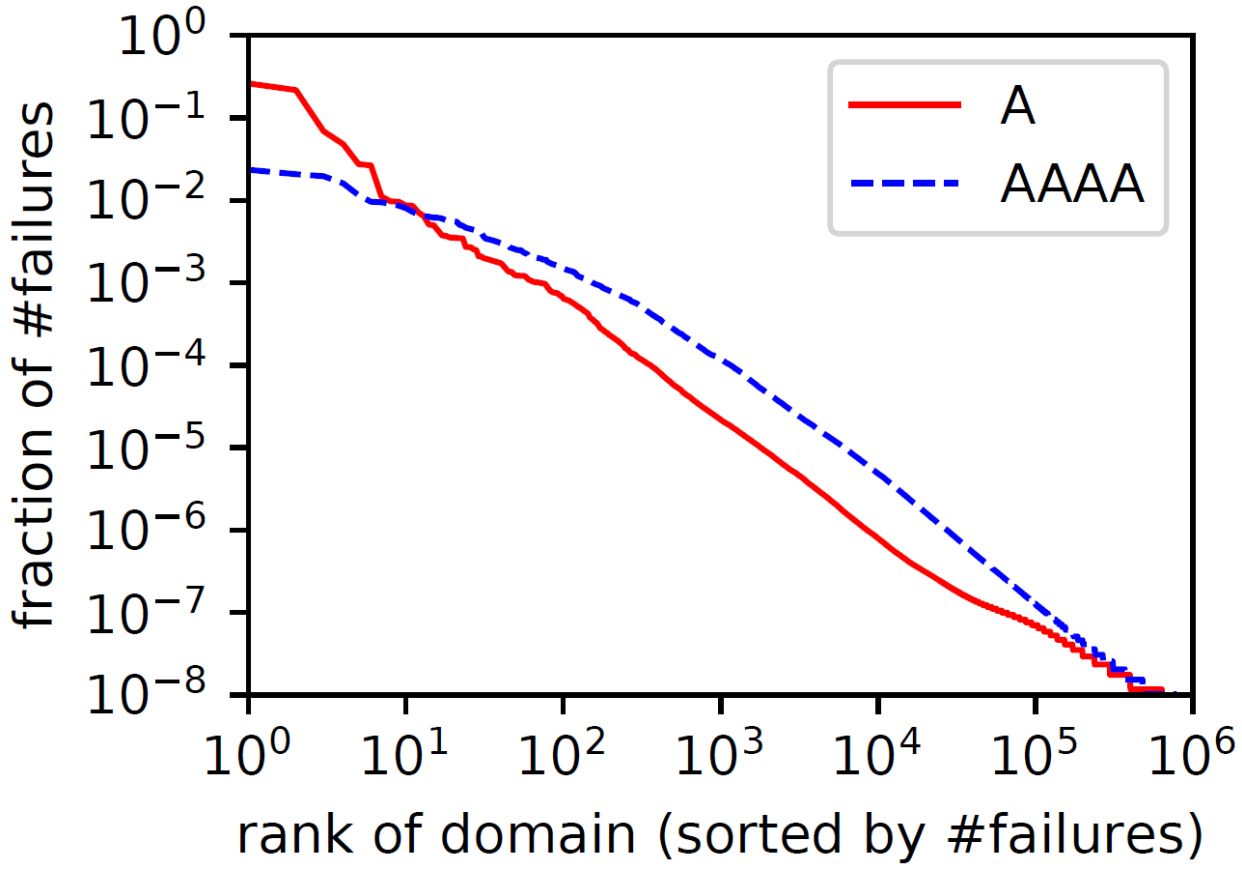

- A queries exhibit high success rates
  - Nevertheless, as many as 7% of domains experience a success rate <50%

# Failures Across Domains



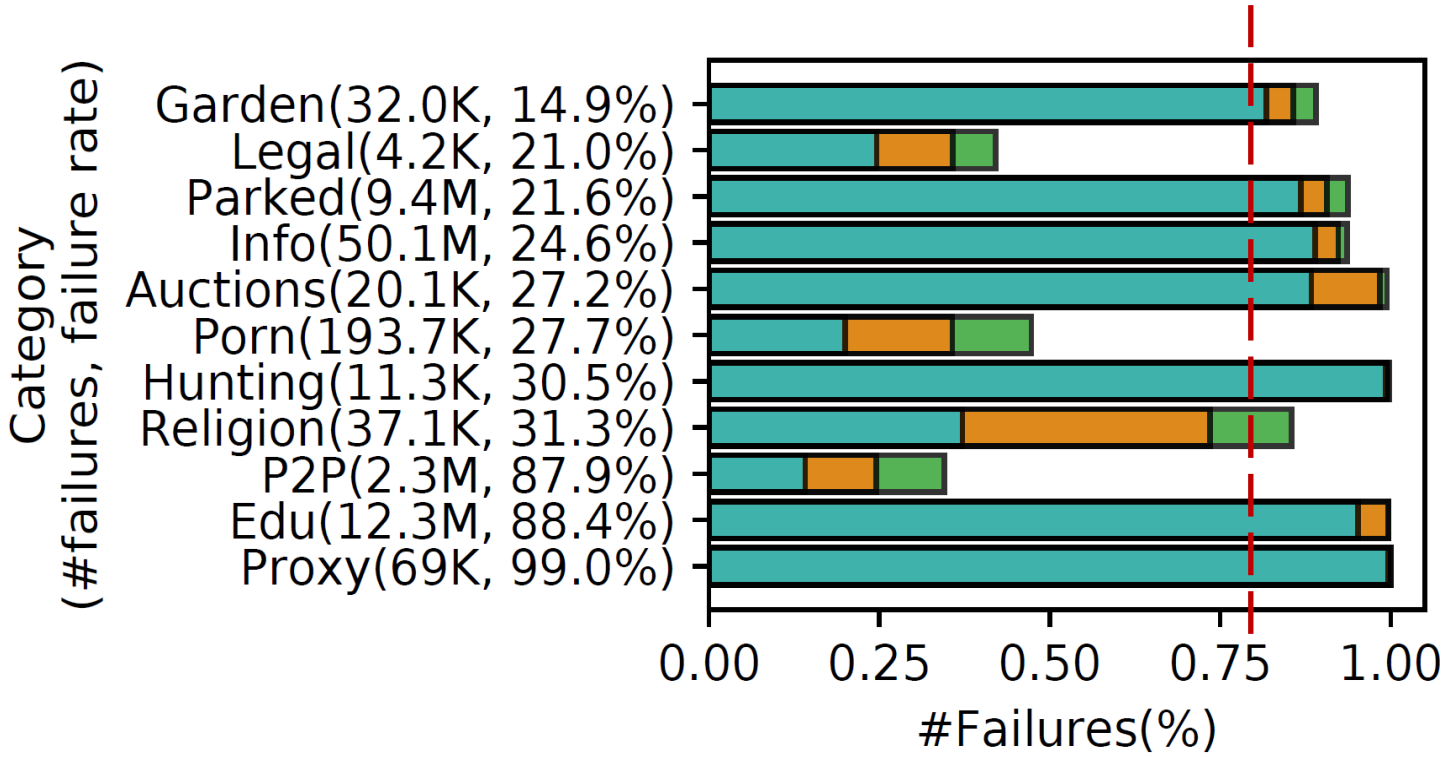

- AAAA queries: ~60% domains have never been successfully resolved
  - Infrastructural limitations in how DNS supports IPv6

# Failures Across Domains



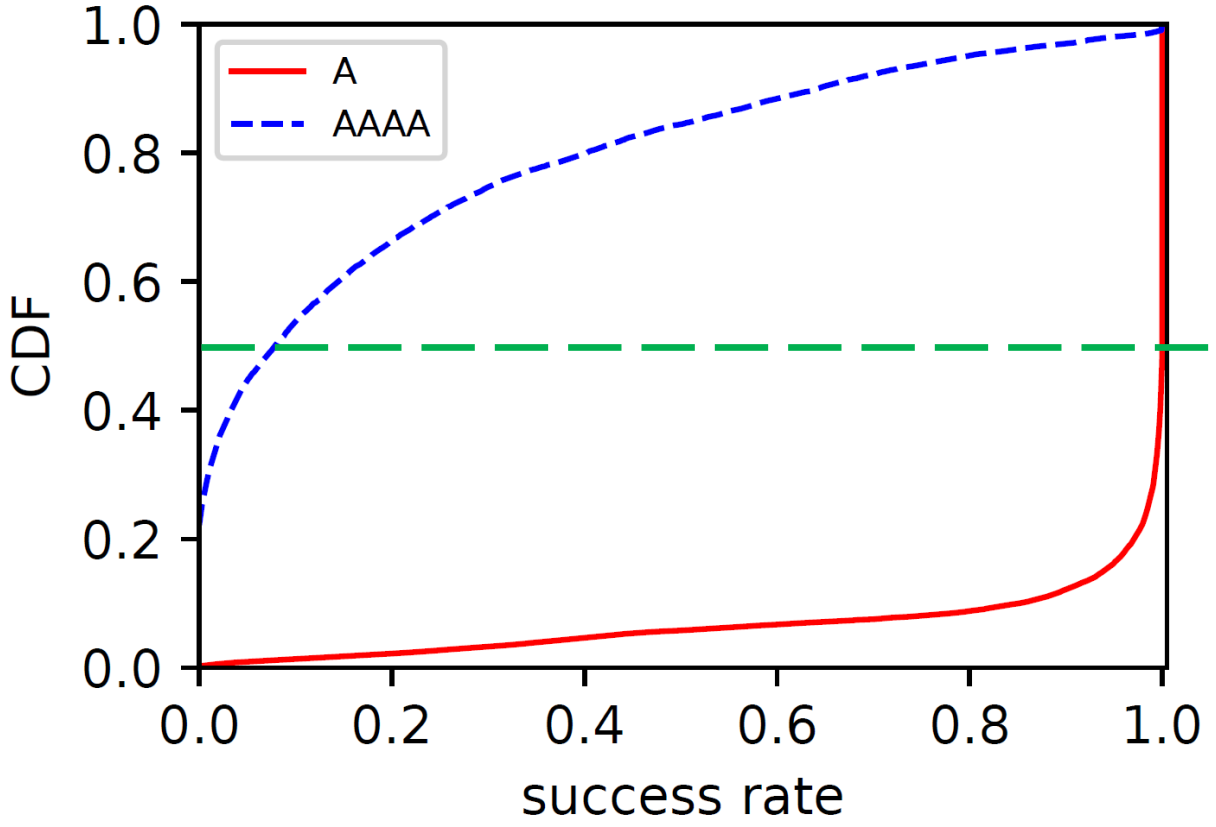

- The concentrate of failures on a small set of domains

# Failures Across Domains



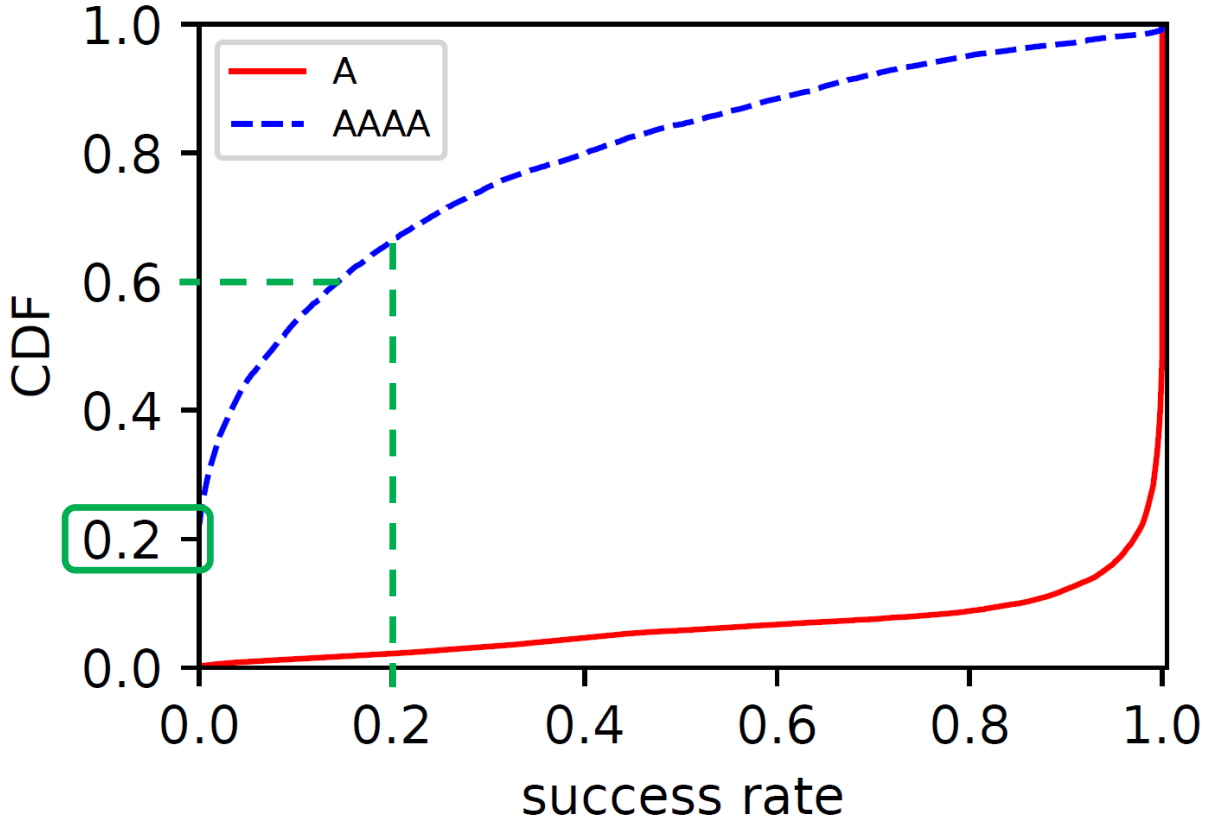

- For most categories, >80% of the failures are attributed to the top 3 SLDs
- Some domain types are paramount in increasing failure rates
  - proxy, porn, parked domains……

# Failures Across Resolvers




- The majority of resolvers serving A queries have very high success rates

# Failures Across Resolvers




- Some resolvers may not be IPv6 ready during our observation period

# Failures Across Resolvers

- Testing public resolvers: #queries (success rate)

| | 114DNS | 360DNS | AlibabaDNS | DNSPOD | GoogleDNS | OpenDNS | ISP | Others |
|---|---|---|---|---|---|---|---|---|
| A | 296.4K(98.5%) | 831.0K(95.9%) | 667.8K(94.7%) | 352.5K(99.6%) | 333.4M(90.7%) | 467.6K(86.3%) | 48.7M(95.3%) | 2.1B(93.5%) |
| AAAA | 75.4K(14.5%) | 50.3K(61.8%) | 112.9K(52.4%) | 15.5K(54.3%) | 40.6M(43.4%) | 31.0K(49.2%) | 9.6M(22.8%) | 252.6M(35.0%) |

# Failures Across Resolvers

- Testing public resolvers: #queries (success rate)

| | 114DNS | 360DNS | AlibabaDNS | DNSPOD | GoogleDNS | OpenDNS | ISP | Others |
|---|---|---|---|---|---|---|---|---|
| A | 296.4K(98.5%) | 831.0K(95.9%) | 667.8K(94.7%) | 352.5K(99.6%) | 333.4M(90.7%) | 467.6K(86.3%) | 48.7M(95.3%) | 2.1B(93.5%) |
| AAAA | 75.4K(14.5%) | 50.3K(61.8%) | 112.9K(52.4%) | 15.5K(54.3%) | 40.6M(43.4%) | 31.0K(49.2%) | 9.6M(22.8%) | 252.6M(35.0%) |

- GoogleDNS dominates the most used public DNS service

# Failures Across Resolvers

- Testing public resolvers: #queries (success rate)

| | 114DNS | 360DNS | AlibabaDNS | DNSPOD | GoogleDNS | OpenDNS | ISP | Others |
|---|---|---|---|---|---|---|---|---|
| A | 296.4K(98.5%) | 831.0K(95.9%) | 667.8K(94.7%) | 352.5K(99.6%) | 333.4M(90.7%) | 467.6K(86.3%) | 48.7M(95.3%) | 2.1B(93.5%) |
| AAAA | 75.4K(14.5%) | 50.3K(61.8%) | 112.9K(52.4%) | 15.5K(54.3%) | 40.6M(43.4%) | 31.0K(49.2%) | 9.6M(22.8%) | 252.6M(35.0%) |

- GoogleDNS dominates the most used public DNS service
- Various success rates: DNSPOD vs OpenDNS

# Failures Across Resolvers

- Testing public resolvers: #queries (success rate)

| | 114DNS | 360DNS | AlibabaDNS | DNSPOD | GoogleDNS | OpenDNS | ISP | Others |
|---|---|---|---|---|---|---|---|---|
| A | 296.4K(98.5%) | 831.0K(95.9%) | 667.8K(94.7%) | 352.5K(99.6%) | 333.4M(90.7%) | 467.6K(86.3%) | 48.7M(95.3%) | 2.1B(93.5%) |
| AAAA | 75.4K(14.5%) | 50.3K(61.8%) | 112.9K(52.4%) | 15.5K(54.3%) | 40.6M(43.4%) | 31.0K(49.2%) | 9.6M(22.8%) | 252.6M(35.0%) |

- GoogleDNS dominates the most used public DNS service
- Various success rates: DNSPOD vs OpenDNS
- AAAA queries: notably lower success rate across all resolvers

# Failures Across Resolvers

- Testing public resolvers: #queries (success rate)

| | 114DNS | 360DNS | AlibabaDNS | DNSPOD | GoogleDNS | OpenDNS | ISP | Others |
|---|---|---|---|---|---|---|---|---|
| A | 296.4K(98.5%) | 831.0K(95.9%) | 667.8K(94.7%) | 352.5K(99.6%) | 333.4M(90.7%) | 467.6K(86.3%) | 48.7M(95.3%) | 2.1B(93.5%) |
| AAAA | 75.4K(14.5%) | 50.3K(61.8%) | 112.9K(52.4%) | 15.5K(54.3%) | 40.6M(43.4%) | 31.0K(49.2%) | 9.6M(22.8%) | 252.6M(35.0%) |

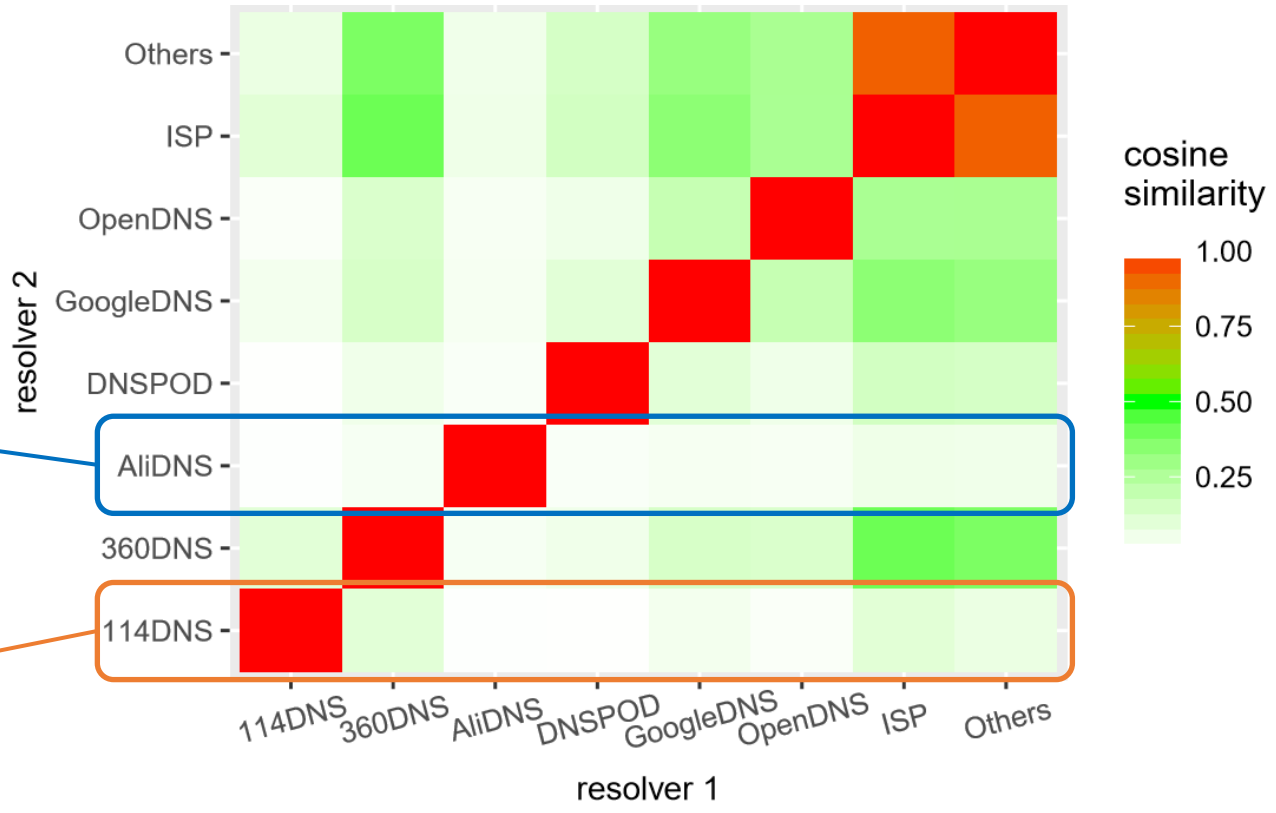
- GoogleDNS dominates the most used public DNS service
- Various success rates: DNSPOD vs OpenDNS
- AAAA queries: notably lower success rate across all resolvers
- Why do public DNS resolvers differ in success rates?

# Failures Across Resolvers

- Comparing domains received between each pair of resolver
- Low similarity with each other
  - Different request patterns

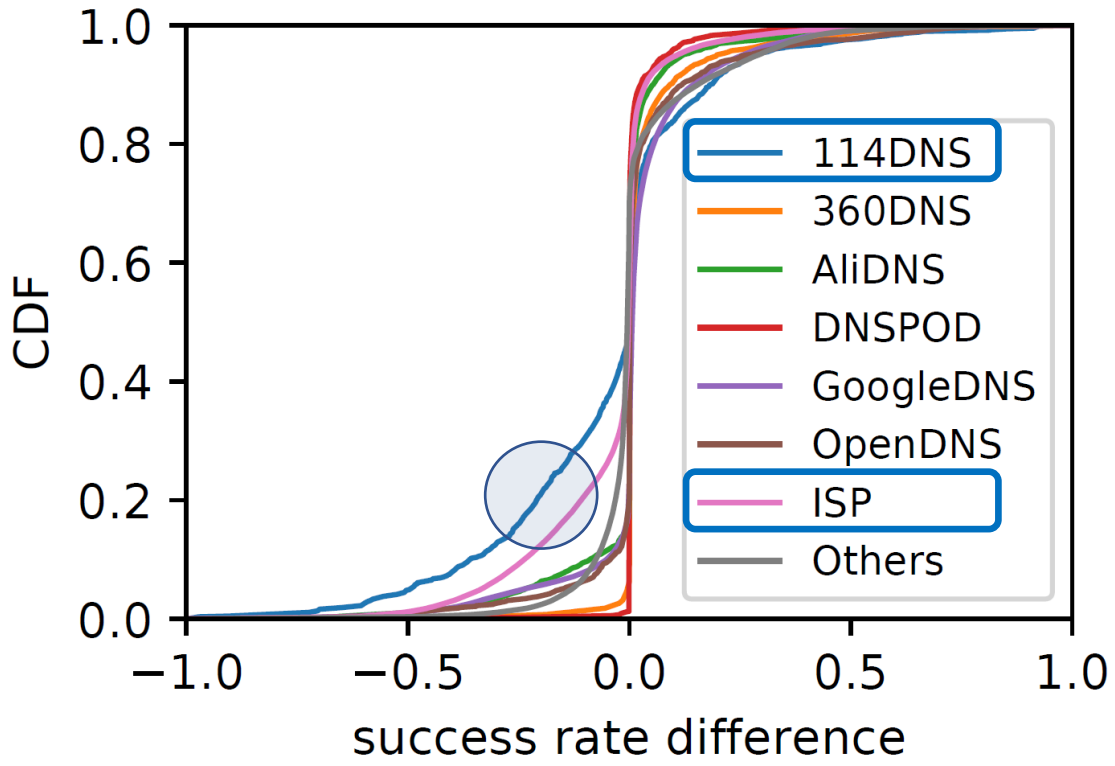
**AliDNS**
- *taobao.com*
- *alipay.com*

**114DNS**
- *akadns.net*
- *akamaiedge.net*



# Failures Across Resolvers

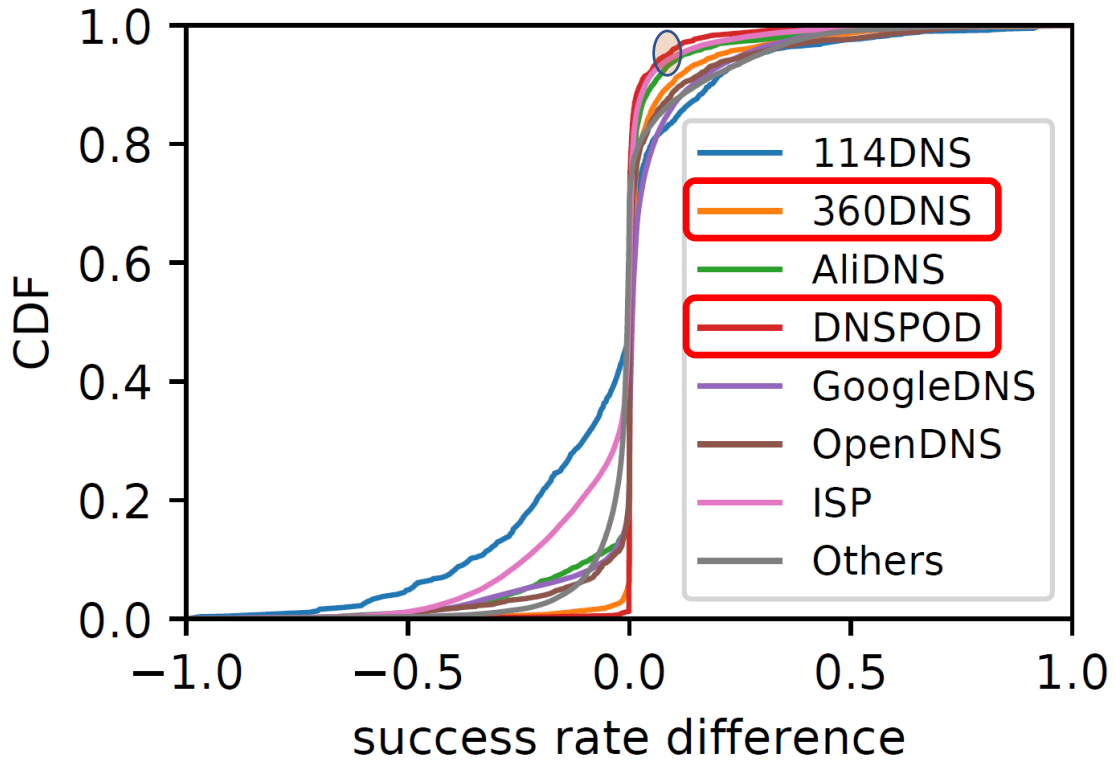
- Comparing infrastructures
  - Compare the success rates of the same domains handled by different resolvers




- Domains resolved by 114DNS and ISP are most likely to fail

# Failures Across Resolvers

- Comparing infrastructures
  - Compare the success rates of the same domains handled by different resolvers




- DNSPOD and 360DNS have higher success rates

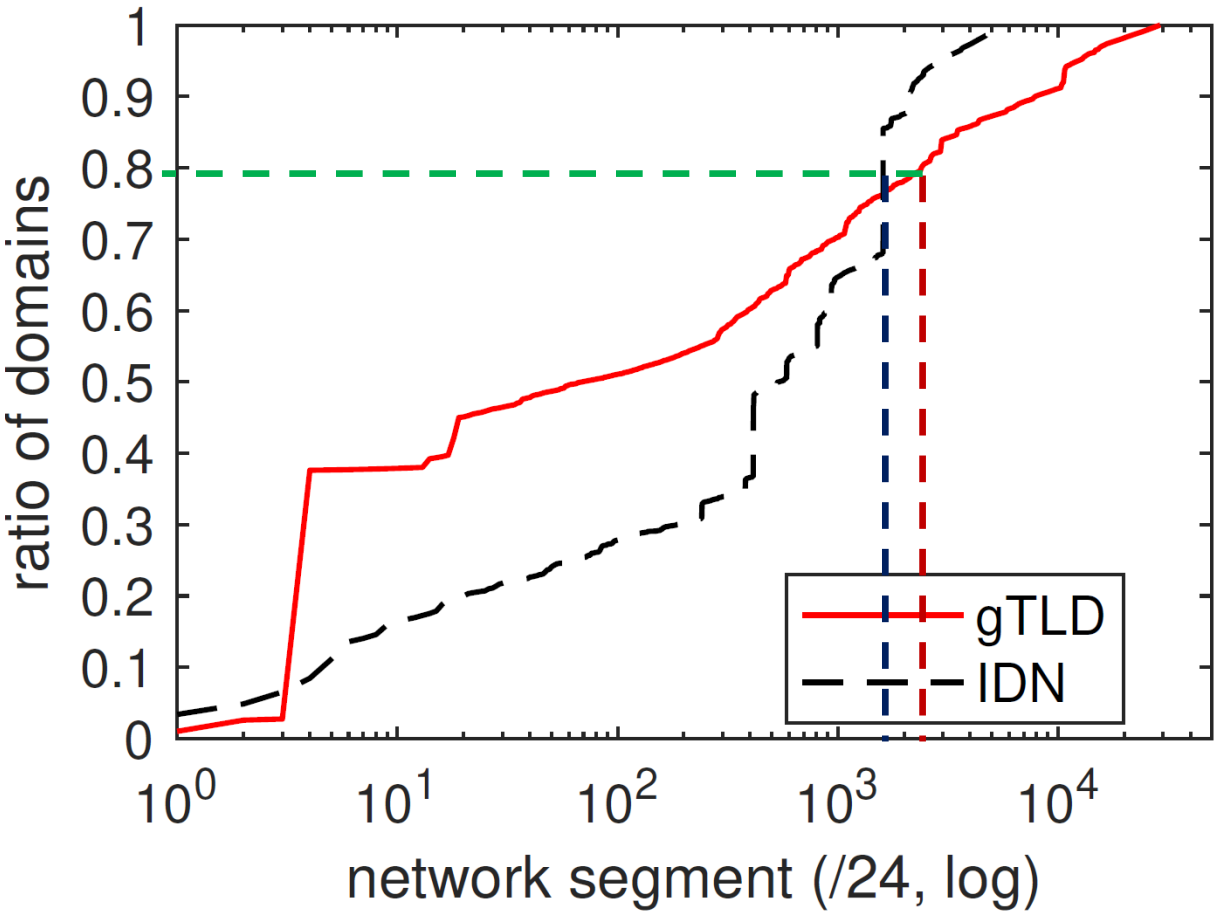
# Failures Across TLDs

- Specifically explore two camps of TLDs
  - The new generic Top Level Domains
  - Those that have Internationalized Domain Name

| | new gTLD | IDN |
|---|---|---|
| total | 4.0M (79.3%) | 0.26M (66.6%) |
| A | 3.4M (88.6%) | 0.17M (86.7%) |
| AAAA | 0.6M (25.9%) | 0.09M (26.4%) |

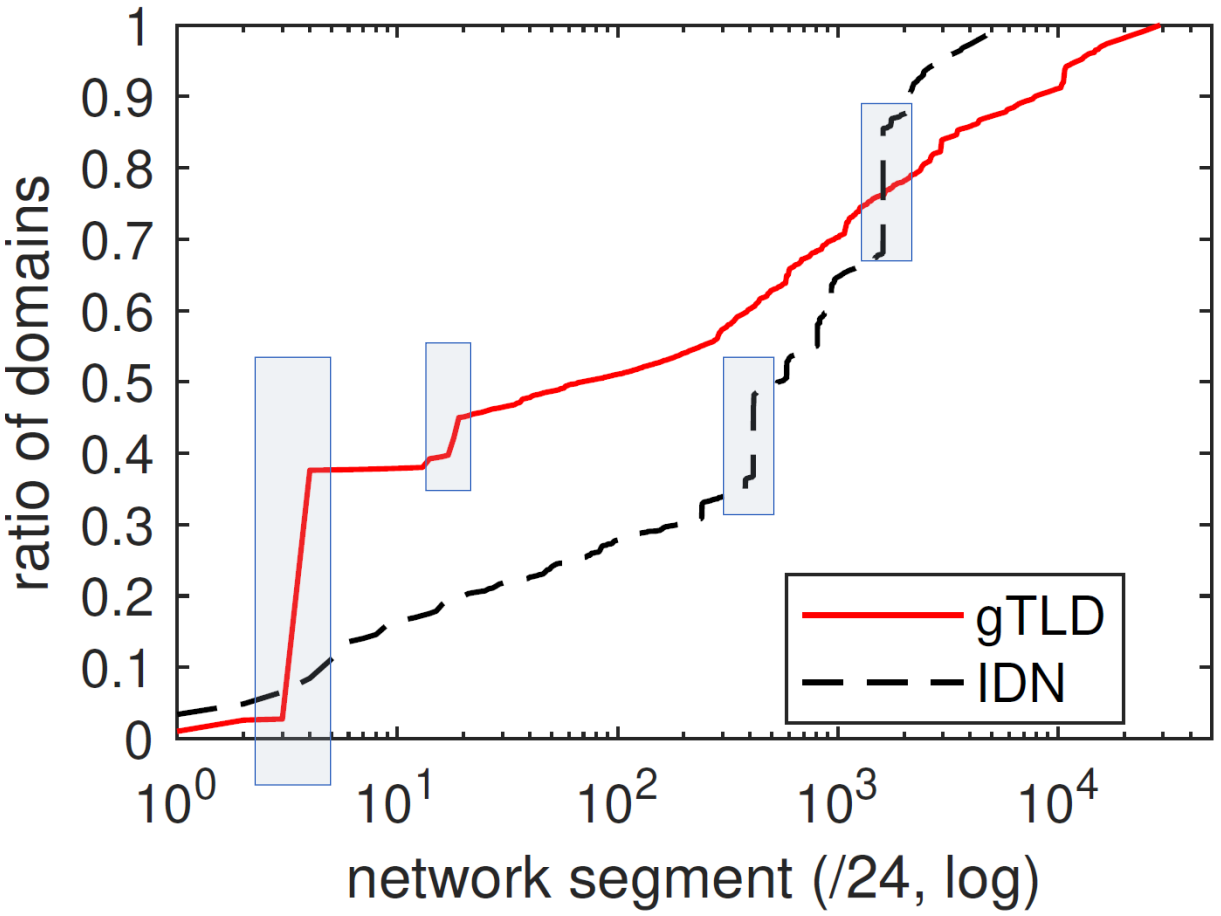
- They show lower success rates, maybe because
  - Such gTLDs attract certain types of domain registrant
  - The presence of malicious domains which are unreliable

# Failures Across TLDs




- The majority of domains map to a relatively small set of prefixes

# Failures Across TLDs




- some /24 network segments serve a large number of domains

# Failures Across TLDs

| No. | subnet | AS num. | AS name | #IPs | #queries | #FQDN | #SLD | #resolvable |
|---|---|---|---|---|---|---|---|---|
| 1 | 23.245.136.0/24 | 18978 | Enzu Inc | 252 | 201.9K (157.1K) | 195.9K (152.2K) | 483 (386) | 0(0) |
| 2 | 192.238.167.0/24 | 395954 | Leaseweb | 236 | 17.4K (14.8K) | 16.3K (13.9K) | 287 (243) | 0 (0) |
| 3 | 172.246.207.0/24 | 18978 | Enzu Inc | 236 | 15.7K (15.4K) | 13.2K (13.0K) | 443(434) | 1 (1) |
| 4 | 104.217.93.0/24 | 40676 | Psychz Net | 253 | 9.0K (1) | 8.8K (1) | 923 (1) | 9(0) |
| 5 | 47.89.58.0/24 | 45102 | Alibaba | 4 | 10.9K (469) | 8.8K (114) | 7.7K (107) | 748 (7) |

# Failures Across TLDs

| No. | subnet | AS num. | AS name | #IPs | #queries | #FQDN | #SLD | #resolvable |
|---|---|---|---|---|---|---|---|---|
| 1 | 23.245.136.0/24 | 18978 | Enzu Inc | 252 | 201.9K (157.1K) | 195.9K (152.2K) | 483 (386) | 0(0) |
| 2 | 192.238.167.0/24 | 395954 | Leaseweb | 236 | 17.4K (14.8K) | 16.3K (13.9K) | 287 (243) | 0 (0) |
| 3 | 172.246.207.0/24 | 18978 | Enzu Inc | 236 | 15.7K (15.4K) | 13.2K (13.0K) | 443(434) | 1 (1) |
| 4 | 104.217.93.0/24 | 40676 | Psychz Net | 253 | 9.0K (1) | 8.8K (1) | 923 (1) | 9(0) |
| 5 | 47.89.58.0/24 | 45102 | Alibaba | 4 | 10.9K (469) | 8.8K (114) | 7.7K (107) | 748 (7) |

- Extremely low rate of successful resolutions today

# Failures Across TLDs

| No. | subnet | AS num. | AS name | #IPs | #queries | #FQDN | #SLD | #resolvable |
|---|---|---|---|---|---|---|---|---|
| 1 | 23.245.136.0/24 | 18978 | Enzu Inc | 252 | 201.9K (157.1K) | 195.9K (152.2K) | 483 (386) | 0(0) |
| 2 | 192.238.167.0/24 | 395954 | Leaseweb | 236 | 17.4K (14.8K) | 16.3K (13.9K) | 287 (243) | 0 (0) |
| 3 | 172.246.207.0/24 | 18978 | Enzu Inc | 236 | 15.7K (15.4K) | 13.2K (13.0K) | 443(434) | 1 (1) |
| 4 | 104.217.93.0/24 | 40676 | Psychz Net | 253 | 9.0K (1) | 8.8K (1) | 923 (1) | 9(0) |
| 5 | 47.89.58.0/24 | 45102 | Alibaba | 4 | 10.9K (469) | 8.8K (114) | 7.7K (107) | 748 (7) |

- The number of queries is close to the number of FQDNs
  - These domains are short-lived and change frequently

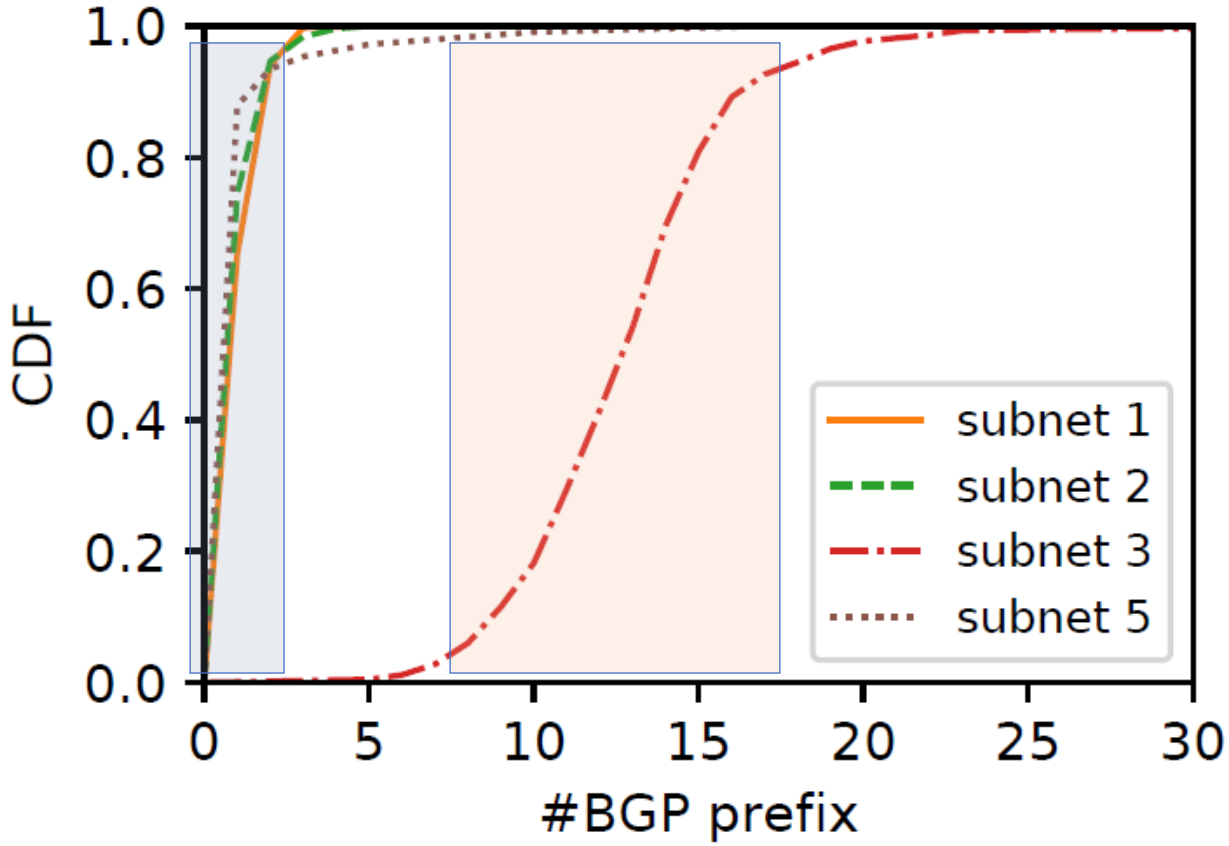
# Failures Across TLDs

| No. | subnet | AS num. | AS name | #IPs | #queries | #FQDN | #SLD | #resolvable |
|---|---|---|---|---|---|---|---|---|
| 1 | 23.245.136.0/24 | 18978 | Enzu Inc | 252 | 201.9K (157.1K) | 195.9K (152.2K) | 483 (386) | 0 (0) |
| 2 | 192.238.167.0/24 | 395954 | Leaseweb | 236 | 17.4K (14.8K) | 16.3K (13.9K) | 287 (243) | 0 (0) |
| 3 | 172.246.207.0/24 | 18978 | Enzu Inc | 236 | 15.7K (15.4K) | 13.2K (13.0K) | 443 (434) | 1 (1) |
| 4 | 104.217.93.0/24 | 40676 | Psychz Net | 253 | 9.0K (1) | 8.8K (1) | 923 (1) | 9 (0) |
| 5 | 47.89.58.0/24 | 45102 | Alibaba | 4 | 10.9K (469) | 8.8K (114) | 7.7K (107) | 748 (7) |

Corresponding to domains classified as malicious

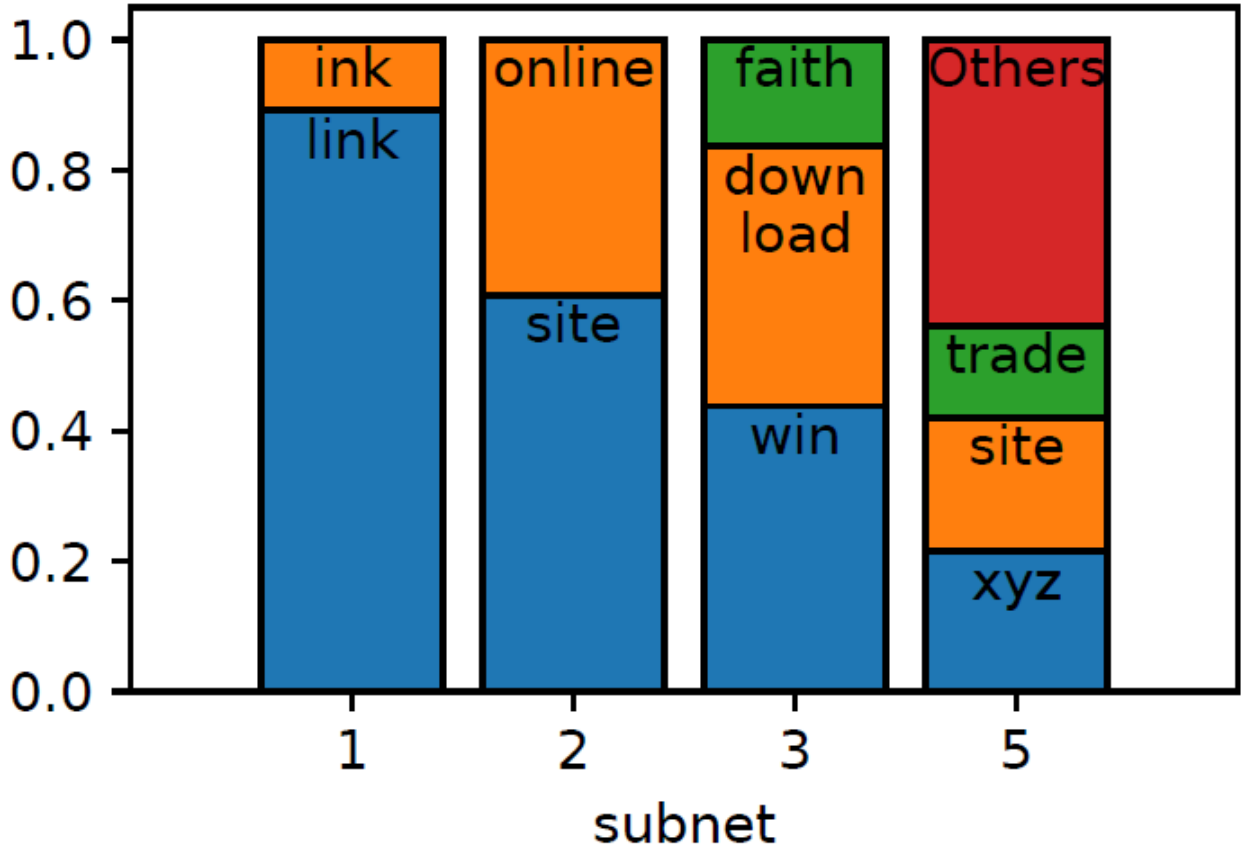
- Two blacklists from VirusTotal and Qihoo 360
- Label a domain as malicious if any of the two blacklists classify it as so

# Failures Across TLDs




- Malicious SLDs hosted in subnet 3 have a larger impact

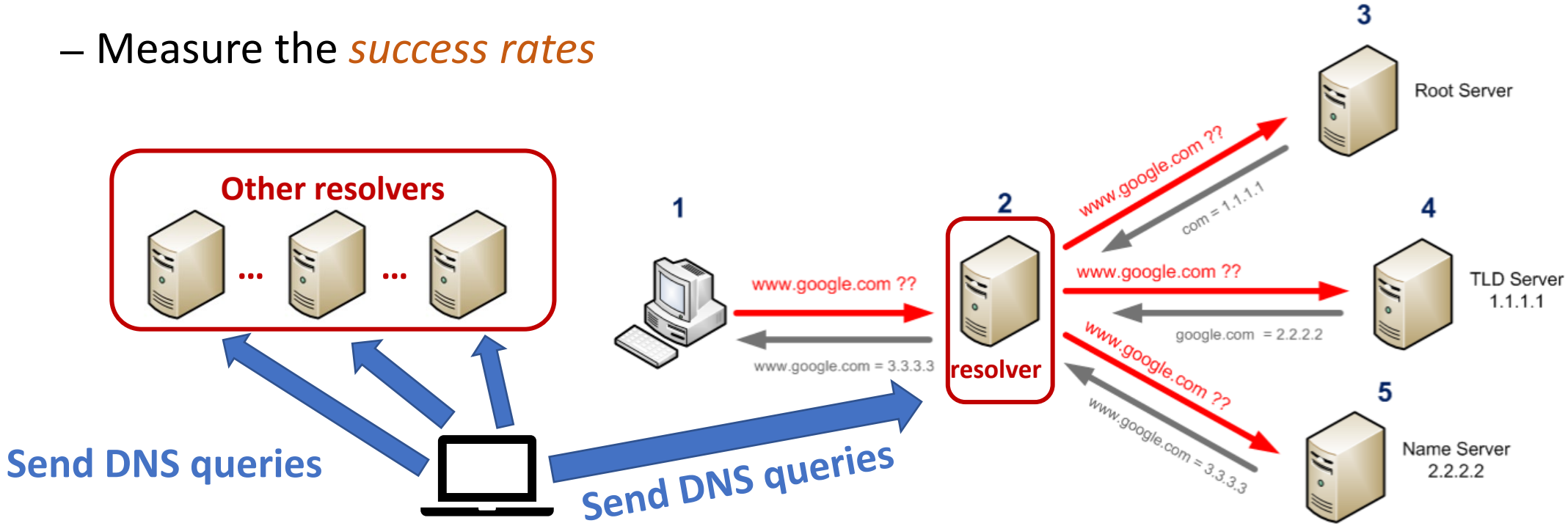
# Failures Across TLDs




- The subnets host different sites mapping to different TLDs

# Implications on Systems Design

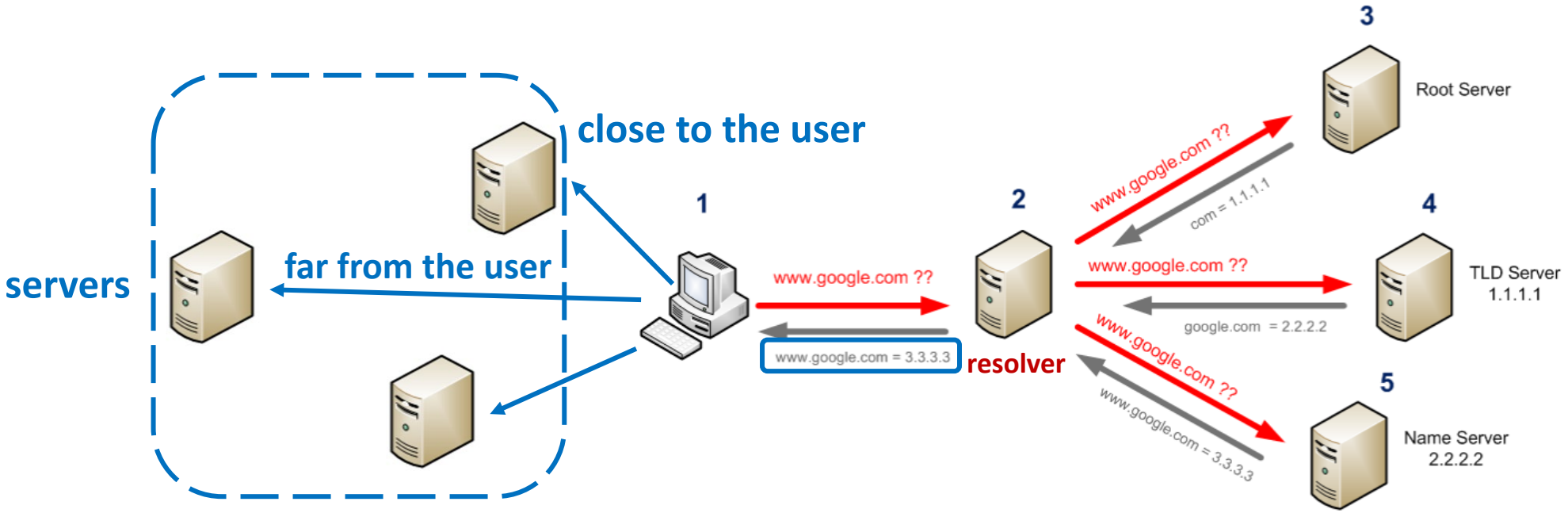
- Active measurement system
  - Distinguish between *resolvers* that support and do not support AAAA queries
  - Test whether a *domain* supports AAAA queries
  - Measure the *success rates*



# Implications on Systems Design

- Active measurement system
  - Localization performance

| 114DNS | 360DNS | AlibabaDNS | DNSPOD | GoogleDNS | OpenDNS | ISP | Others |
|---|---|---|---|---|---|---|---|
| 91.2% | 98.0% | 95.9% | 94.3% | 64.6% | 43.8% | 71.7% | 69.9% |



# Implications on Systems Design

- Such an *active measurement system* is useful for content publishers, ISPs and end users
- For publishers
  - help locate their content
- For ISPs
  - help estimate the IPv6 traffic
- For users
  - help to choose more suitable resolvers

# Implications on Systems Design

| Length | 3 | 4 | 5 | ≥6 |
|---|---|---|---|---|
| % of SLDs | 0.1% | 93.0% | 6.1% | 0.8% |

- Extracting features from domain names may *not* work well for detecting malicious new gTLD domains
- To build a *malicious new gTLD domain detection system*, we could use features like
  - DNS query frequency
  - the number of FQDNs of an SLD
  - the resolved IP addresses
  - the corresponding ASes

# Conclusion

- Findings: based on analysis using passive DNS logs covering over 3B queries from 3 ISPs in China
  - A small number of domains are responsible for the majority of failures
  - Domains and resolvers need to be upgraded for better IPv6 support
  - Diverse failure rates across the DNS resolvers
  - New gTLDs have higher failure rates largely because of malicious domains
- Implications: we propose two potential systems that could build on our findings
  - Active measurement system
  - Malicious new gTLD domain detection system

# Thank you!