Abstract & Introduction

- 本文的工作主要是基于CA/Browser论坛2011年发布的条例对互联网证书进行的评估。
- 作者收集了大量的互联网证书，评估了它们对于条例的遵守情况，以及条例发布前后违规情况的变化趋势。
- 同时作者通过自动化生成概述模板来描述各issuer的证书组成，结合违规统计数据来分析全世界CA的签发证书情况（以此来监视PKI及主动发现重大违规现象）。

Guidelines and Requirements

PKI的安全威胁促使各种规定条例出台，如”Baseline Requirements for the Issuance and Management of Policy-Trusted Certificates”，RFC 5280等，这篇文章主要关注能通过证书审查验证的：Subject identity，Certificate contents，Certificate Extensions，cryptographic algorithm， key requirements。

- Identity Verification and Contents：由于非EV证书的信任级无法确定，条例要求证书中的issuer，Subject，issuance process要有清晰定义。
    - 条例要求叶子证书的最长有效期为5年，2015年4月后降为39个月，EV证书不超过27个月。

TABLE I. X.500 NAME REQUIREMENTS.

X.500 Issuer Fields	
Organization	Required; a name or trademark that identifies the issuing CA
Country	Required; code of country where the CA business is located
Common Name	Optional; if present, should accurately identify the issuing CA

X.500 Subject Fields	
Common Name	Deprecated, must contain a single IP or FQDN if present Subject Alternative Name extension must list applicable names
Organization	Optional, may only appear if verified by the CA Required for extended validation certificates
Location	Covers the Street Address, Locality, State and Postal Code fields Must appear if an Organization name is listed, mustn't otherwise Location must be verified by the CA if present
Country	Required if an organization is listed, must match its location If no organization is listed, may appear based on - the top-level domain of one of the applicable domain name; - IP geolocation of either an applicable IP or the applicant
Registration	Covers Business Category, Incorporation Locality/State/Country Required for extended validation, may not appear otherwise Registration number must also appear in Serial Number field

- Cryptographic Requirements：
  - CA/Browser研讨会接受使用RSA、DSA、EC密钥的证书。
    - RSA密钥至少2048-bit，三个例外允许1024-bit的密钥：（1）2014年之前过期的叶子证书；（2）2011年签发的中间CA证书；（3）2011年之前签发的根证书（只签发叶子证书的根证书）。此外，（1）CA必须保证模数的因子不得小于752，不能是某个素数的幂，不能是已知有漏洞的；（2）e必须是$[2^{16}+1, 2^{256}-1]$区间的一个奇数。
    - DSA的密钥至少2048-bit同时有一个224或256-bit的因子。同时CA要检查生成元的阶。
    - 支持的椭圆曲线为NIST P-256, P-384, P-521。CA必须部分或完全使用NIST SP 800-56A中描

述的ECC公钥校验条例来检查申请人的公钥合法性。

    - 支持的摘要算法为SHA-1,SHA-256, SHA-384, SHA-512。2011年之前签发的根证书可以用MD5自签名。
    - 证书序列号要是非线性的，同时包含至少20-bit的熵。
- Certificate Extensions：对于不同类型的证书（root, intermediate CA or endpoint）条例对Extension的要求不同。

TABLE II. EXTENSIONS OF ENDPOINT CERTIFICATES.

Extension	Requirements
Certificate Policies	Must appear, should not be critical Must include the OID of the issuer's policy May include link to online CPS on issuer website
CRL Distribution Points	Must appear, should not be critical Must include HTTP URL of issuer's CRL file
Authority Information Access	Must appear, must not be critical Must contain HTTP URL of issuer's OCSP service Should contain HTTP URL of issuer's certificate
Basic Constraints	May appear, must be critical if present CA flag must be set to false
Key Usage	May appear, should be critical Must not include "Certificate/CRL Signature"
Extended Key Usage	Must appear, may be critical Must include "Client/Server Authentication" May include "Email Protection" Should not include any other value
Subject Alternative Name	Must appear Should not be critical, unless subject is empty Must include subject's Common Name, if present Must only contain DNS names and IP addresses Should not contain local names or IP addresses

TABLE III. EXTENSIONS OF INTERMEDIATE CA CERTIFICATES.

Extension	Requirements
Certificate Policies	Must appear, should not be critical Must include the OID of the CA's issuance policy May include link to online CPS on issuer website
CRL Distribution Points	Must appear, should not be critical Must include HTTP URL of this CA's CRL file
Authority Information Access	Must appear, must not be critical Must contain HTTP URL of issuer's OCSP service Should contain HTTP URL of issuer's certificate
Basic Constraints	Must appear, must be critical CA flag must be set to true Path Length constraint may be set
Key Usage	Must appear, must be critical Must include "Certificate" and "CRL Signature" May include "Digital Signature" for OCSP signing
Name Constraints	May appear, should be critical if present

TABLE IV. EXTENSIONS OF ROOT CA CERTIFICATES.

Extension	Requirements
Basic Constraints	Must appear, must be critical CA flag must be set to true Path Length constraint should not be set
Key Usage	Must appear, must be critical Must include "Certificate" and "CRL Signature" May include "Digital Signature" for OCSP signing
Extended Key Usage	Must not appear

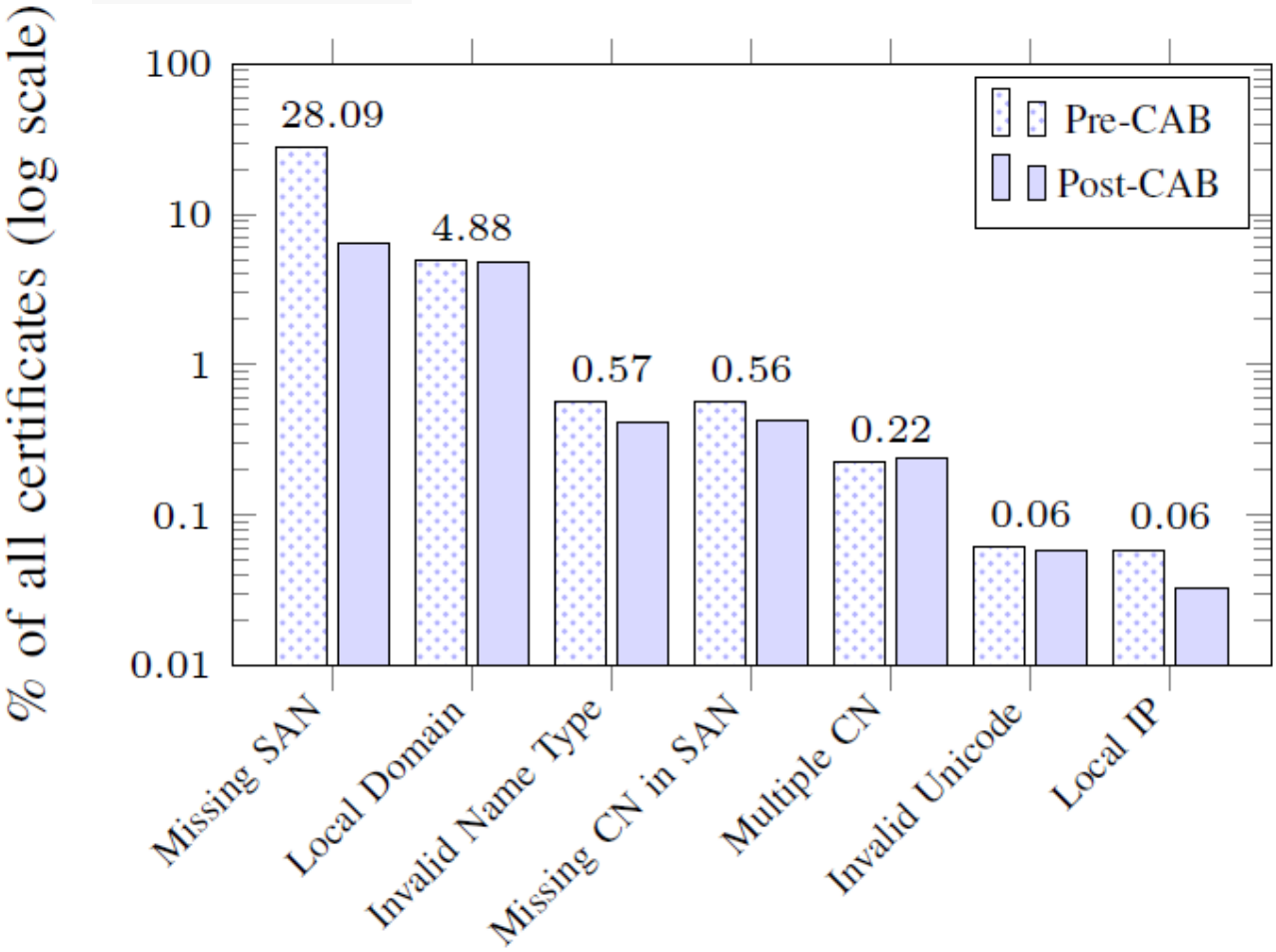
Measuring the Certificate Ecosystem

- Data Collection：从EFF's SSL Observatory的 IP地址和Alexa前100w网站爬取的证书。
  - 共获得了8,349,808个不同的证书；
  - 只关注公共信任的证书（非自签名），及2012前一年及后一年这两年窗口期内的证书，所以一共有1,480,028个。

- Path Reconstruction：这篇文章只关注CA行为，所以不关心证书链是否合法或完整。
    - 主要也是根据subject，issuer，key Identifier等来构建chain。

Global Evaluation

- 根据2012年前、后划分为两个时间段。第一个时间段得到809,425个证书和744个不同的intermediates；第二个时间段得到670,603个证书和668个intermediates。
- 2012之前仅0.39%的证书完全符合基准要求和EV条例，年后上升到0.73%。
- Name Violations：两个时间段平均每个证书代表的合法域名书从1.96涨到2.2，拥有不同二级域名的证书比例上升。52%-56%.



- Issuance and Subject Identity Violations：

- 不同类型证书的签发需求比较稳定：48%,48%,4% VS 49.2%,46.6%,4.2%

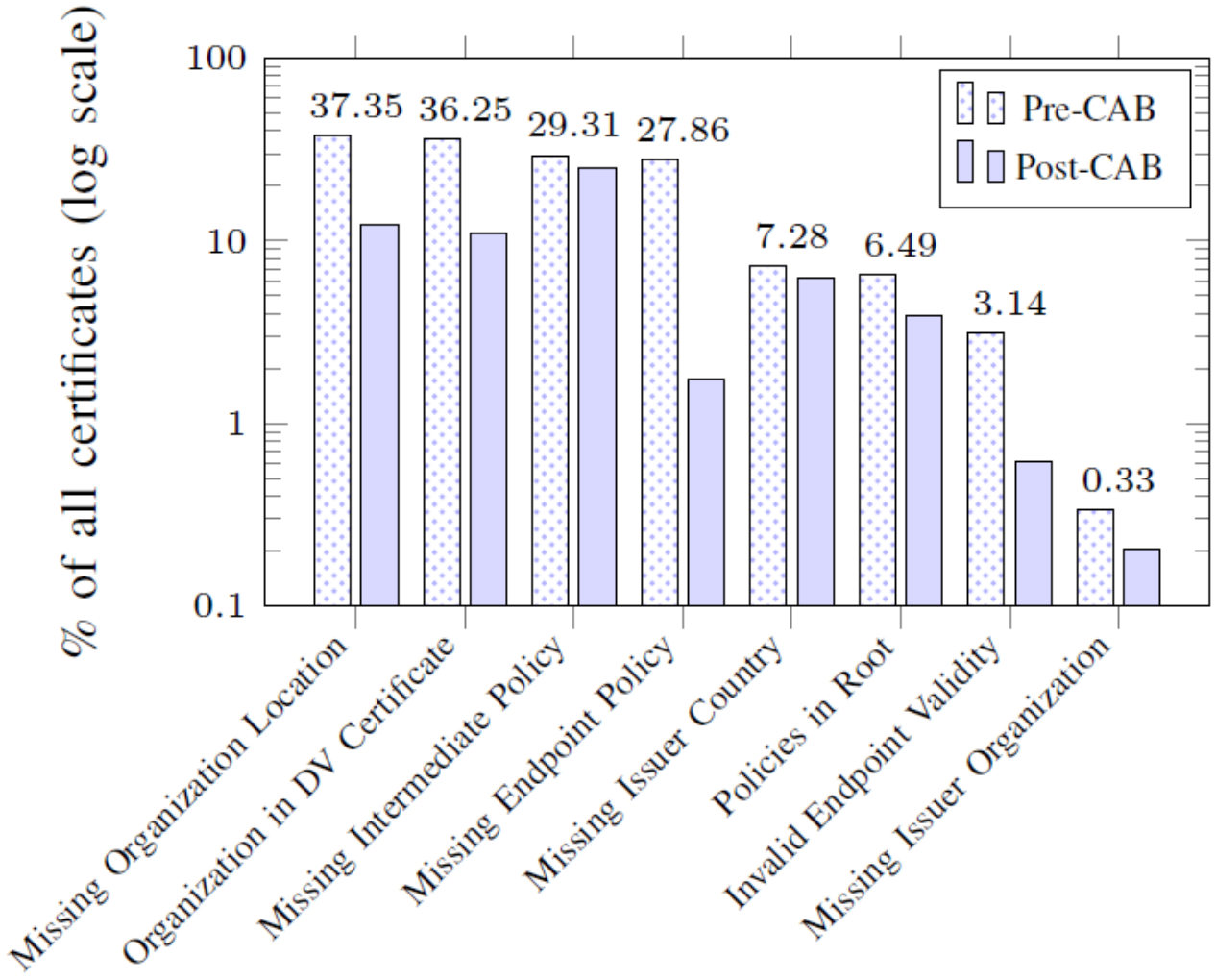


Fig. 3. Identification and Issuance Violations

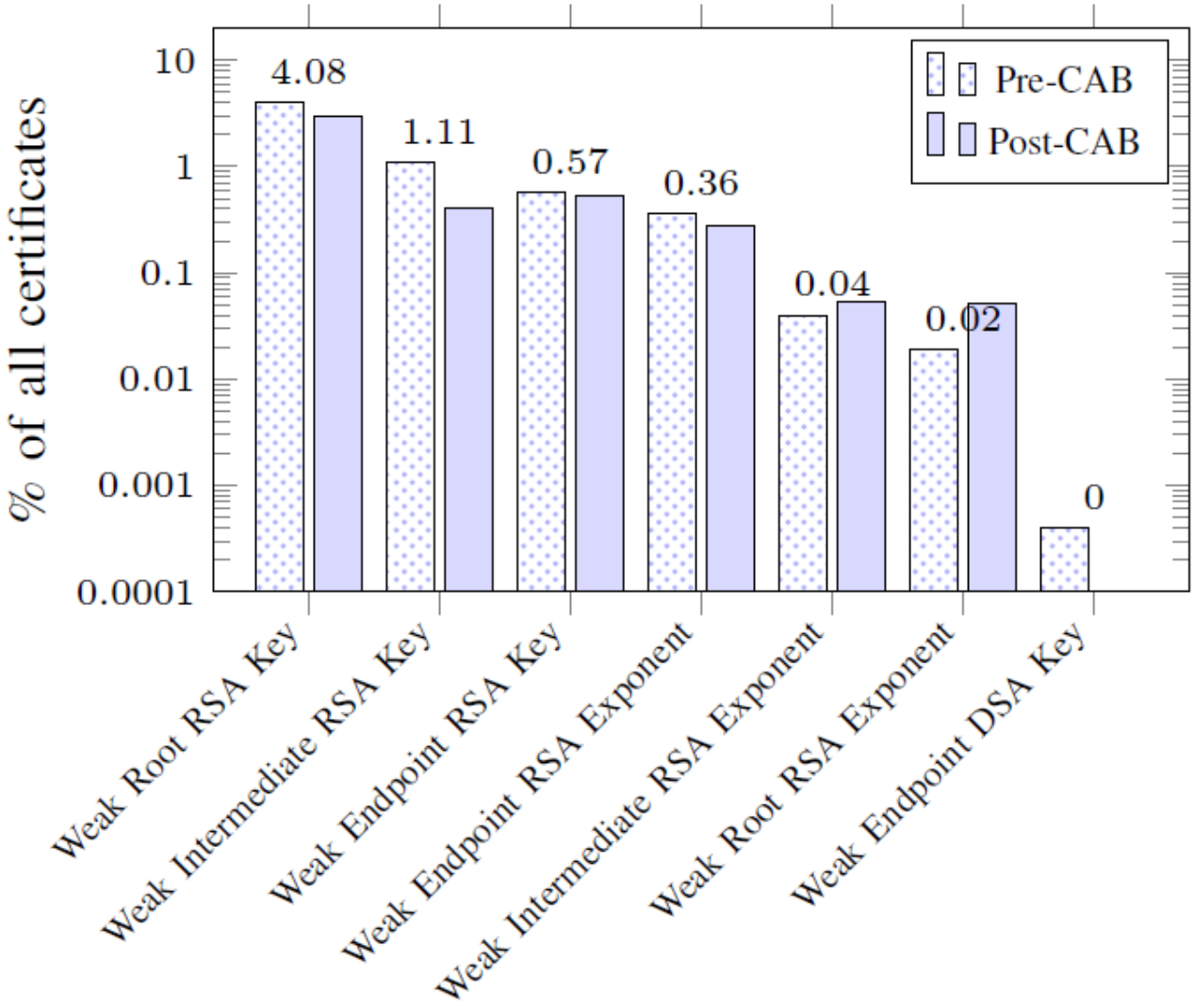


Fig. 5. Cryptographic Violations

- Cryptographic Violations：
  - 大部分都用RSA公钥（3个例外），两个时间段平均模数从1921-bit涨到2017-bit。
  - Web上有一些EC证书，比如Google用的，但只有客户端支持时才用。
  - 2011-2012发现3个DSA证书，两个用了1024-bit的模数，第三个用的512bit的。
  - 支持1024-bit的证书4.3%-5.2%。不过签发1024-bit证书的CA之前发短周期证书。

    - 2011.7.1后没有发现使用MD5签名的证书，2012年之后没有发现证书使用Debian OpenSSL漏洞产生的key。
- Extension Constraints：各类TLS库对扩展约束校验支持不全面。
    - Root Certificates：主要是basic constraints，path length constraints，key usage extension
        - 29.6%的根证书有不合法或不完整basic constraint（说明该证书是否有CA能力），还有6个根证书直接签发端证书，应该根证书离线签发中间证书。
        - 44.7%的根证书不包含key usage extension。（没有这个扩展时证书可以用于任何用途）
        - extended key usage extension可以使证书用于额外用途（如代码签名），2.5%的证书违规。
    - Intermediate CA Certificates：中间CA证书的情况要好很多，所有情况都是因为扩展没被标记为critical。
        - 中间证书不应该拥有给整个互联网空间签名的权利，但只有11个证书用name constraints来约束它们的能力，13年3月后还剩7个。
    - Endpoint Certificates：
        - 最严重的问题是部分端证书的CA-bit，意味着这些端证书可以当CA用。共1.4%的证书，均签发与2012年之前。
        - 好消息是这些证书的中间CA都设置path length为0，坏消息是GnuTLS 3.0之前的版本不检查path length。

        - 不推荐端证书用additional extended key usage，但有2064个证书被用来代码签名，3917个证书包含“Any Key Usage”OID。
    - Revocation Violations：改进良多。
        - OCSP支持率79%-98.7%.
        - 不能检查吊销状态的证书数439-176，来自13个不同的签发者。
    - Path Reconstruction Violations：subject key Identifier（SKI），Authority key Identifier（AKI）和Authority information access（AIA）的支持比率均上升了，可以帮助重建证书链。

Template-level Analysis

- Challenge 1：之前的统计不能反映每个证书签发者的行为。
- Challenge 2：部分根CA会签发一个中间CA证书，然后委托第三方去签发证书。比如Verizon的GTE CyberTrust Global Root，签发了至少40个中间证书，37个都是其他组织在管理。
- 条例里面的很多要求其实是针对CA的而不是单个证书的，于是提出这种基于模板的分析方法。
- Template Clustering：几乎所有CA都使用签发配置文件来签发证书，它包含证书格式、序列号熵源、X.500 subject name中的域，有效期，签名算法，扩展集等，通常这些信息会出现在CA的Certificate Policy Statement（CPS）里。不同的用途和校验方法就对应不同的配置文件签发证书。由于CPS文件通常不是机器可读的，作者试图对证书进行聚类来重构签发配置文件再分析。

    - 把签发过程类似的证书分组，人工介入以消除机器不可读的情况；
    - 将不同群集里的违规情况进行对比，以发现不同CA及委托第三方的差异性。
- The Clustering Algorithm：
    - 使用一个基于特征向量的距离度量来判断证书的相似性（相似方式签发的）。相关特征有的是数字的（有效期），有的是分类的（签名算法），有的是属性集（扩展）。数字特性用L1度量，分类的用离散方法度量（d(x,y) = 1 iff x = y, 否则是0），属性集用Jaccard距离度量。
    - The clustering procedure applies the k-medoid algorithm seeded with the k-means++ initialization step.
- Cluster Evaluation：
    - 检查集群中心，记录违规情况；
    - 用一个规则集来检验单个证书，以判断集群的质量和集群的相关性。
    - 对每个集群出现的template-specific违规情况，检查证书域名的合法性和对应的IP位置。

TABLE V. CLUSTERING FEATURES.

High Weight	Medium Weight	Low Weight
Parent CA Signature and key algorithms Set of X.509 extensions Policy identifiers Authority information access Key usage, basic constraints	Subject name fields CRL distribution points Extended key usage	Key size Issuance date Validity period Serial number length

Clustering Results and Discussions

- 一个可视化工具：产生了571个群集，每个至少包含5个证书。

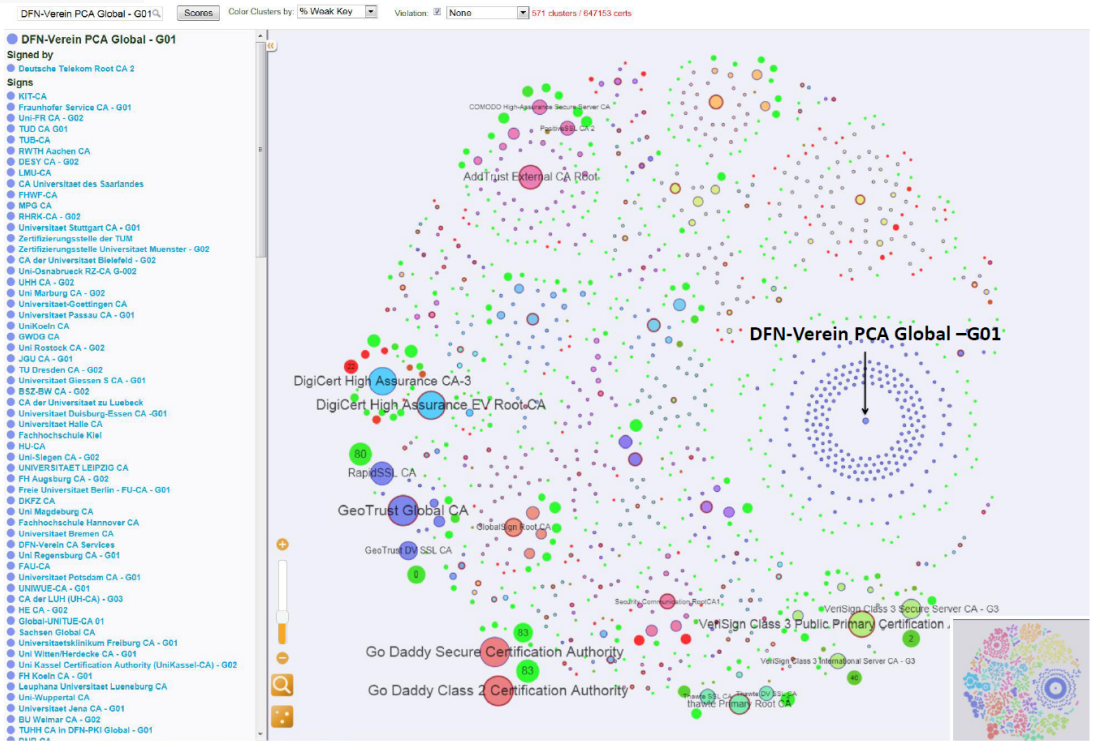


Fig. 11. Distribution of Clusters among CAs. The color scheme reflect the percentage of weak keys in a cluster. The left pane shows the searching interface.

- CA规模与遵守条例的关系：
  - 长尾效应。
  - 委托越多违规越多，小CA违规情况多。
  - CA数量增长速度降低。
  - 非常合规的CA也容易有几个违规多的模板。

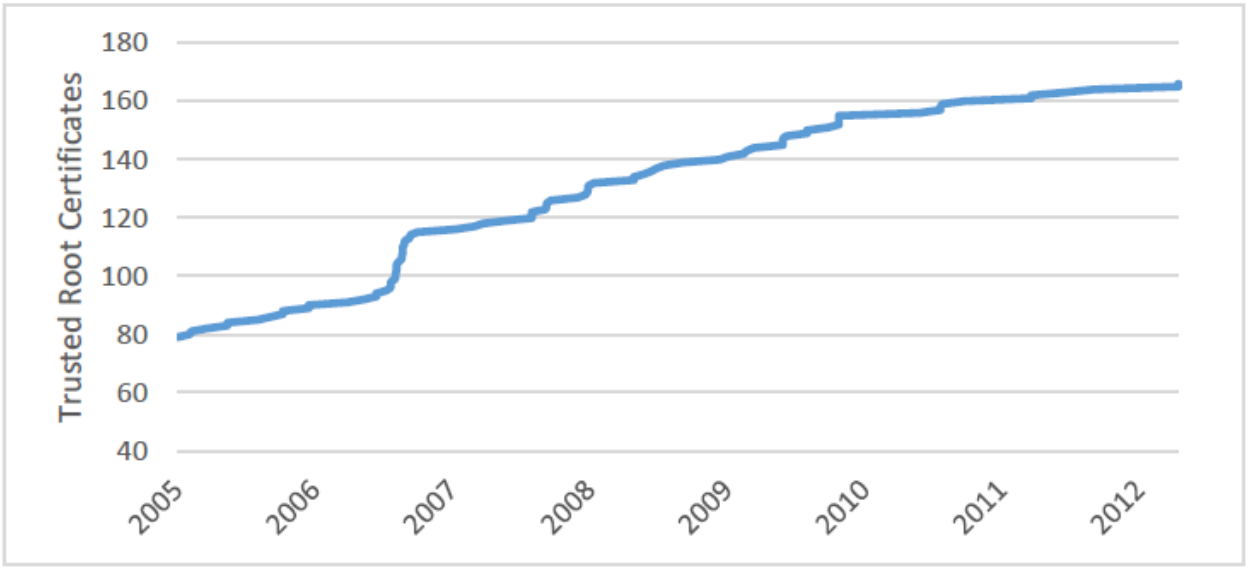


Fig. 15. Growth of the Mozilla Root Program.

- DNS分析：解析域名的IP地址，分析服务器的位置是不是和证书列出的国家匹配。发现少量域名所有权更替这类的情况。
- CDN网络：老问题。