

Imperfect Forward Secrecy How Diffie-Hellman Fails in Practice

David Adrian, Karthikeyan Bhargavan, Zakir Durumeric,
Pierrick Gaudry, Matthew Green, J. Alex Halderman,
Nadia Heninger, Drew Springall, Emmanuel Thomé,
Luke Valenta, Benjamin VanderSloot, Eric Wustrow,
Santiago Zanella-Béguelin, Paul Zimmermann

weakdh.org

July 7, 2015



New Directions in Cryptography

Invited Paper

WHITFIELD DIFFIE AND MARTIN E. HELLMAN, MEMBER, IEEE

Textbook Diffie-Hellman

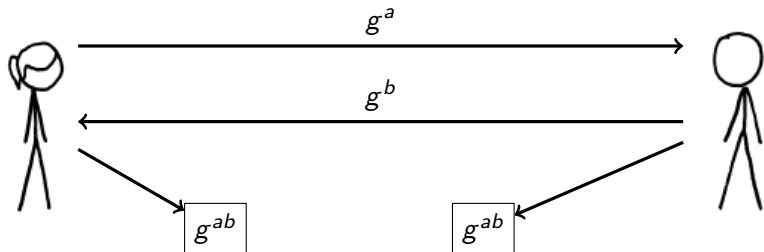
[Diffie Hellman 1976]

Public Parameters

G a group (e.g. \mathbb{F}_p^*)

g group generator

Key Exchange



Reminders: Orders and groups

For our purposes, g generates multiplicative group mod p .

Theorem (Fermat's Little Theorem)

$g^{p-1} \equiv 1 \pmod{p}$ for any $0 < g < p$.

Let $\text{ord}(g)_p$ be the order of $g \pmod{p}$. (Smallest positive integer such that $g^{\text{ord}(g)_p} \equiv 1 \pmod{p}$.)

Theorem (Lagrange)

$\text{ord}(g)_p$ divides $p - 1$.

Discrete Log

Problem: Given $t = g^\ell$, compute ℓ .

Warm-up: Elementary discrete log algorithms

Pollard rho, baby step giant step

- ▶ $O(\sqrt{q})$ for (sub)group of order q .

Warm-up: Elementary discrete log algorithms

Pollard rho, baby step giant step

- ▶ $O(\sqrt{q})$ for (sub)group of order q .

Pollard lambda

- ▶ $O(\sqrt{r})$ for exponent in size range r .

Warm-up: Elementary discrete log algorithms

Pollard rho, baby step giant step

- ▶ $O(\sqrt{q})$ for (sub)group of order q .

Pollard lambda

- ▶ $O(\sqrt{r})$ for exponent in size range r .

Pohlig-Hellman

1. Factor group order $q = \prod_i q_i^{e_i}$.
2. Solve discrete log in each subgroup in time $e_i \sqrt{q_i}$.
3. Use Chinese remainder theorem to reconstruct $\log \bmod q$.

Warm-up: Elementary discrete log algorithms

Pollard rho, baby step giant step

- ▶ $O(\sqrt{q})$ for (sub)group of order q .

Pollard lambda

- ▶ $O(\sqrt{r})$ for exponent in size range r .

Pohlig-Hellman

1. Factor group order $q = \prod_i q_i^{e_i}$.
2. Solve discrete log in each subgroup in time $e_i \sqrt{q_i}$.
3. Use Chinese remainder theorem to reconstruct $\log \bmod q$.

Best practice: g should generate large prime-order subgroup mod p . Common choices include “safe” primes $p = 2q + 1$ or DSA groups.

Pop quiz

Do people deploying cryptography follow best practice?

Diffie-Hellman deployment in practice

- ▶ 23.9% of 14.3M HTTPS servers support Diffie-Hellman

Diffie-Hellman deployment in practice

- ▶ 23.9% of 14.3M HTTPS servers support Diffie-Hellman
- ▶ Observed 70,000 distinct primes p

Diffie-Hellman deployment in practice

- ▶ 23.9% of 14.3M HTTPS servers support Diffie-Hellman
- ▶ Observed 70,000 distinct primes p
- ▶ Found 4800 groups (p, g) where $(p - 1)/2$ was not prime.

Diffie-Hellman deployment in practice

- ▶ 23.9% of 14.3M HTTPS servers support Diffie-Hellman
- ▶ Observed 70,000 distinct primes p
- ▶ Found 4800 groups (p, g) where $(p - 1)/2$ was not prime.
- ▶ Applied ECM to opportunistically factor $(p - 1)/2$.

Diffie-Hellman deployment in practice

- ▶ 23.9% of 14.3M HTTPS servers support Diffie-Hellman
- ▶ Observed 70,000 distinct primes p
- ▶ Found 4800 groups (p, g) where $(p - 1)/2$ was not prime.
- ▶ Applied ECM to opportunistically factor $(p - 1)/2$.
- ▶ Learned prime factors of order of g for 750 groups, used in 40,000 connections across our Internet scans.

Diffie-Hellman deployment in practice

- ▶ 23.9% of 14.3M HTTPS servers support Diffie-Hellman
- ▶ Observed 70,000 distinct primes p
- ▶ Found 4800 groups (p, g) where $(p - 1)/2$ was not prime.
- ▶ Applied ECM to opportunistically factor $(p - 1)/2$.
- ▶ Learned prime factors of order of g for 750 groups, used in 40,000 connections across our Internet scans.

For random p , $p - 1$ likely to have large factors, so might not recover full discrete log.

Short exponents with composite group orders: A sad tale

[van Oorschot, Wiener]

- ▶ Some implementations use short exponents ℓ : 128, 160 bits.

Short exponents with composite group orders: A sad tale

[van Oorschot, Wiener]

- ▶ Some implementations use short exponents ℓ : 128, 160 bits.
- ▶ If $\ell < \prod_i q_i^{e_i}$ for $q_i^{e_i} \mid \text{ord}_p(g)$, Pohlig-Hellman over only these subgroups reconstructs ℓ by CRT.

Short exponents with composite group orders: A sad tale

[van Oorschot, Wiener]

- ▶ Some implementations use short exponents ℓ : 128, 160 bits.
- ▶ If $\ell < \prod_i q_i^{e_i}$ for $q_i^{e_i} \mid \text{ord}_p(g)$, Pohlig-Hellman over only these subgroups reconstructs ℓ by CRT.
- ▶ Using this attack we computed secret exponent for 159 exchanges and partial information in 460 exchanges.

Subexponential algorithms for prime fields: Index calculus

Goal: Solve $g^\ell \equiv t \pmod{p}$.

Definition: An integer is *B-smooth* if all its prime factors are $\leq B$.

Fix some a priori bound B .

Subexponential algorithms for prime fields: Index calculus

Goal: Solve $g^\ell \equiv t \pmod{p}$.

Definition: An integer is *B-smooth* if all its prime factors are $\leq B$.
Fix some a priori bound B .

1. **Relation finding:** Enumerate pairs of B -smooth integers equivalent mod p .

$$p_1^{a_{11}} \dots B^{a_{1k}} = 1 \equiv p + 1 = p_1^{r_{11}} p_2^{r_{12}} \dots B^{r_{1k}}$$

$$p_1^{a_{21}} \dots B^{a_{2k}} = 2 \equiv p + 2 = p_1^{r_{21}} p_2^{r_{22}} \dots B^{r_{2k}}$$

\vdots

$$p_1^{a_{k1}} \dots B^{a_{kk}} = z \equiv p + z = p_1^{r_{k1}} p_2^{r_{k2}} \dots B^{r_{kk}}$$

Index calculus: Linear algebra

Take log of both sides. Assume subgroup of order q . Then

$$a_{11} \log p_1 + \cdots + a_{1k} \log B \equiv r_{11} \log p_1 + \cdots + r_{1k} \log B \pmod{q}$$

$$a_{21} \log p_1 + \cdots + a_{2k} \log B \equiv r_{21} \log p_1 + \cdots + r_{2k} \log B \pmod{q}$$

\vdots

$$a_{k1} \log p_1 + \cdots + a_{kk} \log B \equiv r_{k1} \log p_1 + \cdots + r_{kk} \log B \pmod{q}$$

Also get some relations for free: $\log -1 = (p-1)/2$ etc.

2. **Linear Algebra:** Solve system of equations for $\log p_i$:

$$\log p_1 \equiv s_1$$

\vdots

$$\log p_k \equiv s_k$$

Actually computing individual logs

Input target t .

3. Try to find some B -smooth value

$$g^R t = p_1^{e_1} \dots B^{e_B}$$

Then using known values of $\log p_i$ write

$$\log t = -R + e_1 \log p_1 + \dots + e_B \log B \pmod{q}$$

Index calculus running time

1. **Relation collection** Runtime depends on (1) work to test if integer is B -smooth, (2) probability integer is B -smooth, (3) B .
2. **Linear algebra** Runtime depends on cost of sparse linear algebra for B -dimensional matrix mod q .
3. **Individual log** Runtime depends on probability that $g^R t$ is B -smooth.

Optimizing for B gives runtime of

$$\exp((\sqrt{2} + o(1))\sqrt{\log p \log \log p}) = L_p(1/2, \sqrt{2})$$

Number field sieve

[Gordon], [Joux, Lercier], [Semaev]

1. **Polynomial selection:** Find a polynomial f and an integer m such that $f(m) \equiv 0 \pmod{p}$, $\deg f = 5$ or 6 , coeffs of f relatively small. Defines a number field $\mathbb{Q}(x)/f(x)$.

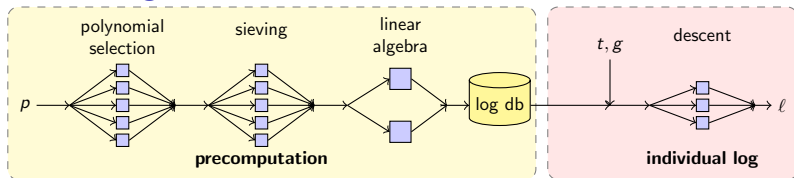
For $\gamma = \sum_i a_i \alpha^i$ in ring of integers, define homomorphism $\varphi(\gamma) = \sum_i a_i m^i$ to $\mathbb{Z}/p\mathbb{Z}$.

2. **Relation collection** Collect relations of form

$$p_1^{a_{11}} \dots \mathfrak{B}^{a_{1k}} = a + b\alpha \equiv a + bm = p_1^{r_{11}} \dots B^{r_{1k}}$$

3. **Linear algebra** Once there are enough relations, solve for $\log p_i$.
4. **Individual log** "Descent" Try to write target t as sum of logs in known database.

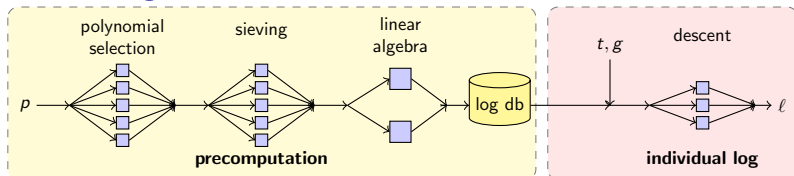
Implementing the NFS with CADO-NFS



$L(1/3, 1.923)$

$L(1/3, 1.232)$

Implementing the NFS with CADO-NFS

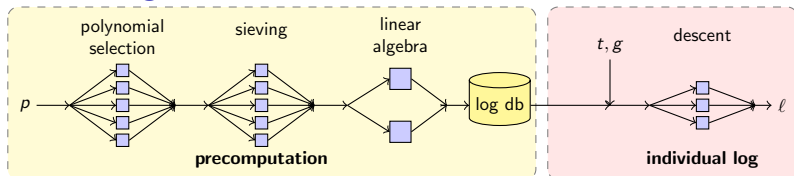


$L(1/3, 1.923)$

$L(1/3, 1.232)$

	Sieving			Linear Algebra		Descent
	I	1pb	core-years	rows	core-years	core-time
RSA-512	14	29	0.5	4.3M	0.33	
DH-512	15	27	2.5	2.1M	7.7	10 mins

Implementing the NFS with CADO-NFS



$L(1/3, 1.923)$

$L(1/3, 1.232)$

	Sieving			Linear Algebra		Descent
	I	1pb	core-years	rows	core-years	core-time
RSA-512	14	29	0.5	4.3M	0.33	
DH-512	15	27	2.5	2.1M	7.7	10 mins

Times for cluster computation:

	polysel	sieving	linalg	descent
	2000-3000 cores		288 cores	24 cores
DH-512	3 hours	15 hours	120 hours	90 seconds

www.eecs.berkeley.edu

Your connection to this site is private.

Permissions

Connection



The identity of this website has been verified by InCommon RSA Server CA but does not have public audit records.

[Certificate Information](#)



Your connection to www.eecs.berkeley.edu is encrypted with obsolete cryptography.

The connection uses TLS 1.2.

The connection is encrypted using AES_256_CBC, with SHA1 for message authentication and ECDHE_RSA as the key exchange mechanism.



Site information

You have never visited this site before today.

[What do these mean?](#)



ole External Relations Calendar

Calendar Highlights

Monday, July 6

Talk: Challenges in P2P Botnet Monitoring

Shankar Karuppayah, TU Darmstadt / CASED
1-2 p.m., 380 Soda Hall

Friday, July 24

An Empirical Understanding of Conflict-Driven Clause-Learning SAT Solvers

Vijay Ganesh, University of Waterloo
12-1 p.m., 373 Soda Hall

Thursday, August 27

Airbnb Info Session/Tech Talk

5-8 p.m., Wozniak Lounge/430/438 Soda Hall

Friday, August 28

Palantir Technologies Info Session/Tech Talk

5-8 p.m., Wozniak Lounge/430/438 Soda Hall

benefit society.

ly 1



Eli Yablonovitch has been awarded the **Isaac Newton Medal** by the Institute of Physics in London. This award is the institute's highest honor and is given to any physicist, regardless of subject area, background or nationality for outstanding contributions to physics. Prof.

TLS Diffie-Hellman Key Exchange

TLS = Transport Layer Security

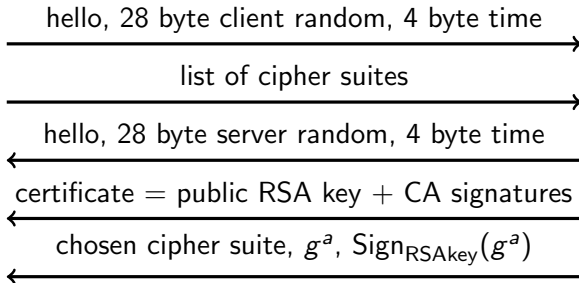
hello, 28 byte client random, 4 byte time

list of cipher suites



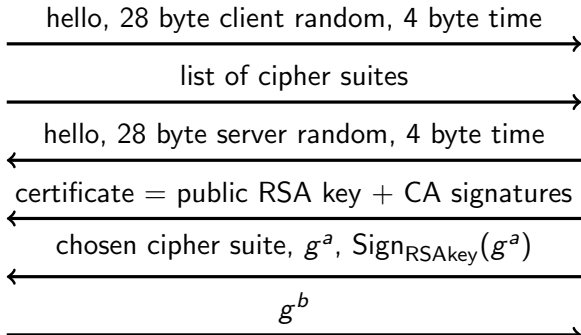
TLS Diffie-Hellman Key Exchange

TLS = Transport Layer Security



TLS Diffie-Hellman Key Exchange

TLS = Transport Layer Security



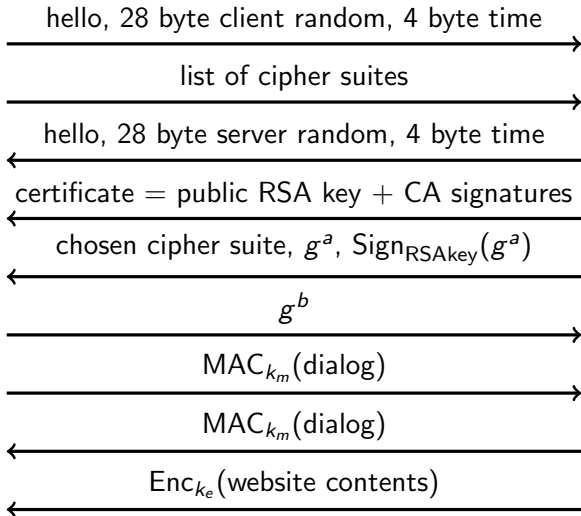
$F(g^{ab},$
randoms)
 $\rightarrow k_m, k_e$



$F(g^{ab},$
randoms)
 $\rightarrow k_m, k_e$

TLS Diffie-Hellman Key Exchange

TLS = Transport Layer Security

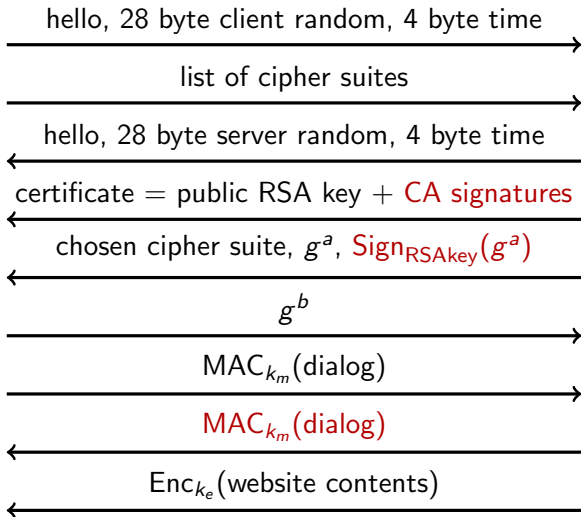


$F(g^{ab},$
randoms)
 $\rightarrow k_m, k_e$

$F(g^{ab},$
randoms)
 $\rightarrow k_m, k_e$



Authentication in TLS Diffie-Hellman Key Exchange

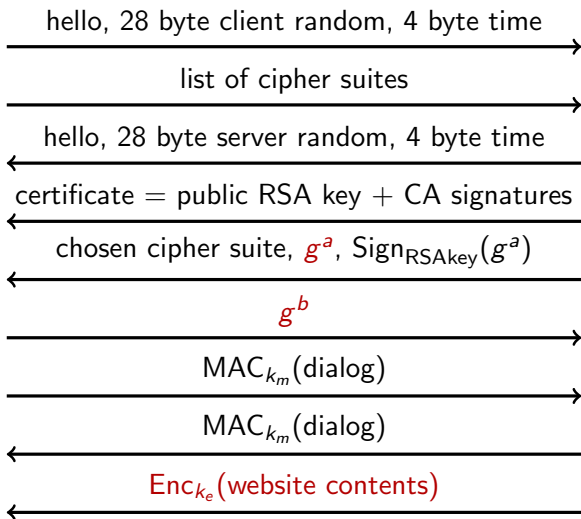


$F(g^{ab},$
randoms)
→ k_m, k_e



$F(g^{ab},$
randoms)
→ k_m, k_e

Encryption in TLS Diffie-Hellman Key Exchange



$F(g^{ab},$
randoms)
→ k_m, k_e

$F(g^{ab},$
randoms)
→ k_m, k_e



Some differences between practice and theory

- ▶ Cipher-suite and parameter negotiation.
- ▶ Politics.
- ▶ ElGamal is uncommon.
- ▶ Group negotiation.
- ▶ Short exponents.
- ▶ Long-tailed distribution of group parameters.
- ▶ Primes with many bits in common.

Does anyone *use* 512-bit Diffie-Hellman?

International Traffic in Arms Regulations

April 1, 1992 version

Category XIII--Auxiliary Military Equipment ...

(b) Information Security Systems and equipment, cryptographic devices, software, and components specifically designed or modified therefore, including:

(1) Cryptographic (including key management) systems, equipment, assemblies, modules, integrated circuits, components or software with the capability of maintaining secrecy or confidentiality of information or information systems, except cryptographic equipment and software as follows:

(i) Restricted to decryption functions specifically designed to allow the execution of copy protected software, provided the decryption functions are not user-accessible.

(ii) Specially designed, developed or modified for use in machines for banking or money transactions, and restricted to use only in such transactions. Machines for banking or money transactions include automatic teller machines, self-service statement printers, point of sale terminals or equipment for the encryption of interbanking transactions.

...

Commerce Control List: Category 5 - Info. Security

(From 2014)

a.1.a. A symmetric algorithm employing a key length in excess of 56-bits; or

a.1.b. An asymmetric algorithm where the security of the algorithm is based on any of the following:

a.1.b.1. Factorization of integers in excess of 512 bits (e.g., RSA);

a.1.b.2. Computation of discrete logarithms in a multiplicative group of a finite field of size greater than 512 bits (e.g., Diffie-Hellman over Z/pZ); or

a.1.b.3. Discrete logarithms in a group other than mentioned in 5A002.a.1.b.2 in excess of 112 bits (e.g., Diffie-Hellman over an elliptic curve);

a.2. Designed or modified to perform cryptanalytic functions;

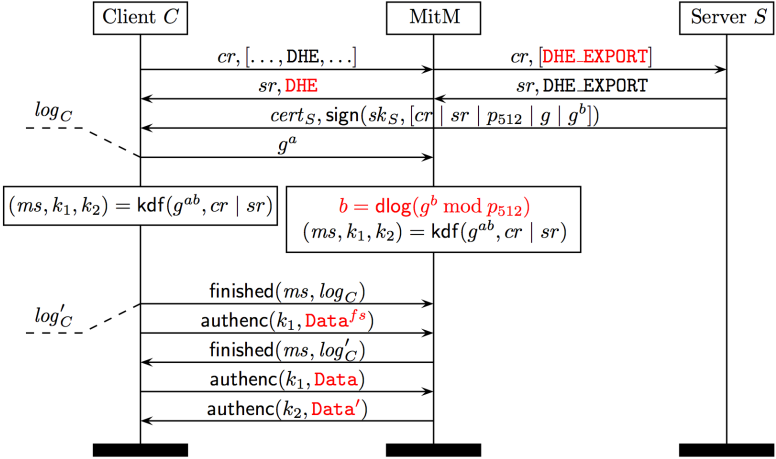
Export cipher suites in TLS

```
TLS_RSA_EXPORT_WITH_RC4_40_MD5  
TLS_RSA_EXPORT_WITH_RC2_CBC_40_MD5  
TLS_RSA_EXPORT_WITH_DES40_CBC_SHA  
TLS_DH_RSA_EXPORT_WITH_DES40_CBC_SHA  
TLS_DHE_DSS_EXPORT_WITH_DES40_CBC_SHA  
TLS_DHE_RSA_EXPORT_WITH_DES40_CBC_SHA  
TLS_DH_Annon_EXPORT_WITH_RC4_40_MD5  
TLS_DH_Annon_EXPORT_WITH_DES40_CBC_SHA
```

In March 2015, export cipher suites supported by 36.7% of the 14 million sites serving browser-trusted certificates!

FREAK attack [BDFKPSZZ 2015]: Use fast 512-bit factorization to downgrade modern browsers to broken export-grade RSA.

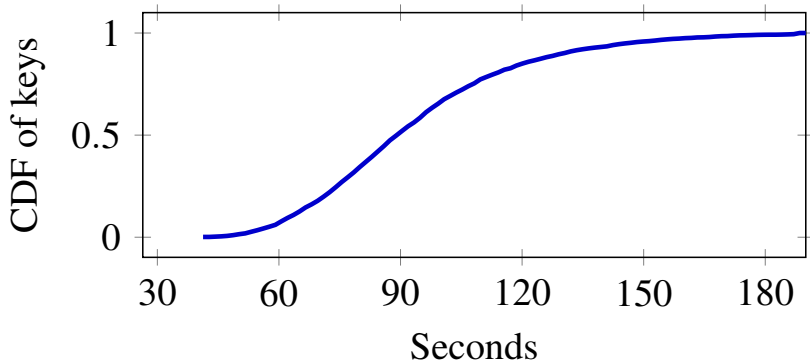
Man-in-the-middle downgrade attack to export DH



Protocol flaw: Server does not sign chosen cipher suite.

Carrying out the downgrade attack

- ▶ In April 2015, 8.4% of Alexa Top 1M HTTPS domains supported DHE_EXPORT.
- ▶ 82% use most common prime. 10% use 2nd most common prime.
- ▶ We carried out precomputation for these primes. \approx 1 week each on 2000-3000 cores.
- ▶ Individual descent times on 24-core machine:



Many hosts use the same group parameters.

What about passive attacks for 768 or 1024-bit keys?

Cost estimates for discrete log

	Sieving			Linear Algebra		Descent
	I	l _{pb}	core-years	rows	core-years	core-time
RSA-512	14	29	0.5	4.3M	0.33	
DH-512	15	27	2.5	2.1M	7.7	10 mins
RSA-768	16	37	800	250M	100	
DH-768	17	35	8,000	150M	28,500	2 days
RSA-1024	18	42	1,000,000	8.7B	120,000	
DH-1024	19	40	10,000,000	5.2B	35,000,000	30 days

- ▶ Use RSA-768 factoring record, adjust to sieve less and modern processors.
- ▶ For discrete log, oversieve to decrease linear algebra cost.
- ▶ For descent, used early-abort ECM implementation to get experimental timings.

Cost estimates for discrete log

	Sieving			Linear Algebra		Descent
	I	1pb	core-years	rows	core-years	core-time
RSA-512	14	29	0.5	4.3M	0.33	
DH-512	15	27	2.5	2.1M	7.7	10 mins
RSA-768	16	37	800	250M	100	
DH-768	17	35	8,000	150M	28,500	2 days
RSA-1024	18	42	1,000,000	8.7B	120,000	
DH-1024	19	40	10,000,000	5.2B	35,000,000	30 days

- ▶ [Geiselman Steinwandt] give ASIC sieving design. 10x their estimates at modern sizes = \$8M to sieve 1024-bit DL in one year.
- ▶ Titan supercomputer = 300,000 cores → 117 years for linear algebra.
- ▶ Assuming 80x speedup for ASIC linear algebra = hundreds of millions of \$.

James Bamford, 2012, Wired

According to another top official also involved with the program, the NSA made an enormous breakthrough several years ago in its ability to cryptanalyze, or break, unfathomably complex encryption systems employed by not only governments around the world but also many average computer users in the US. The upshot, according to this official: “Everybody’s a target; everybody with communication is a target.”

[...]

The breakthrough was enormous, says the former official, and soon afterward the agency pulled the shade down tight on the project, even within the intelligence community and Congress. “Only the chairman and vice chairman and the two staff directors of each intelligence committee were told about it,” he says. The reason? “They were thinking that this computing breakthrough was going to give them the ability to crack current public encryption.”

2013 NSA “Black Budget”

“Also, we are investing in groundbreaking cryptanalytic capabilities to defeat adversarial cryptography and exploit internet traffic.”



4. Communicate Results

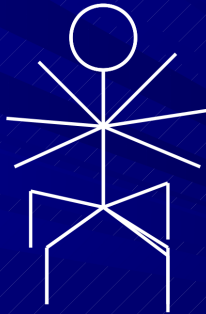


Can we decrypt the VPN traffic?

- If the answer is “No” then explain how to turn it into a “YES!”
- If the answer is “YES!” then...



Happy Dance!!



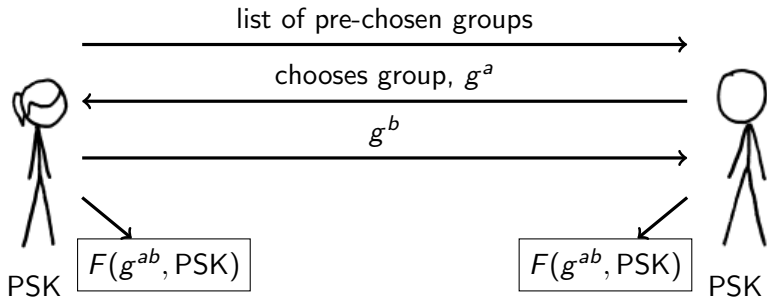


Turn that Frown Upside Down! From “No” to “YES!”

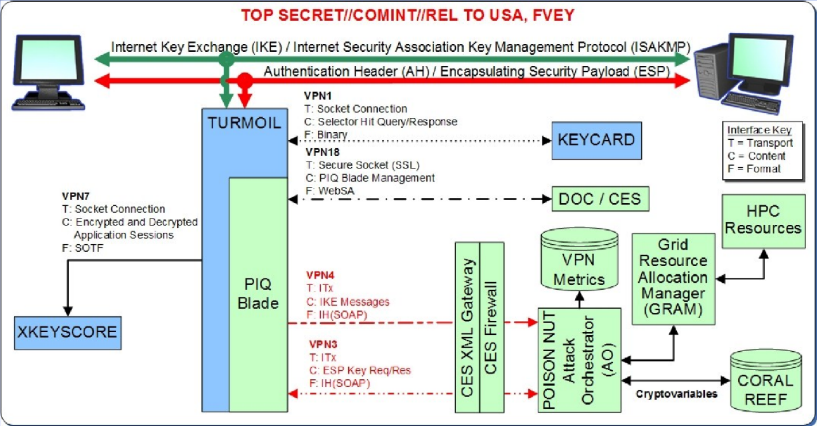


- Depends on why we couldn't decrypt it
- Find Pre-Shared Key
- Locate complete paired collect
- Locate both IKE and ESP traffic
- Have collection sites do surveys for the IP's
- Find better quality collect with rich metadata

IKE Key Exchange for VPNs/IPsec



NSA VPN Attack Orchestration



Vulnerable servers, if the attacker can precompute for . . .

	all 512-bit p	all 768-bit p	one 1024-bit p	ten 1024-bit p
HTTPS Top 1M MITM	45K (8.4%)	45K (8.4%)	205K (37.1%)	309K (56.1%)
HTTPS Top 1M	118 (0.0%)	407 (0.1%)	98.5K (17.9%)	132K (24.0%)
HTTPS Trusted MITM	489K (3.4%)	556K (3.9%)	1.84M (12.8%)	3.41M (23.8%)
HTTPS Trusted	1K (0.0%)	46.7K (0.3%)	939K (6.56%)	1.43M (10.0%)
IKEv1 IPv4	–	64K (2.6%)	1.69M (66.1%)	1.69M (66.1%)
IKEv2 IPv4	–	66K (5.8%)	726K (63.9%)	726K (63.9%)
SSH IPv4	–	–	3.6M (25.7%)	3.6M (25.7%)

Practical Mitigations

- ▶ Move to elliptic curve cryptography
- ▶ If ECC isn't an option, move to ≥ 2018 -bit primes.
- ▶ If 2048-bit primes aren't an option, generate a fresh 1024-bit prime.
- ▶ Major browsers will refuse to accept 512-bit groups, will sunset 768 and 1024-bit soon.

Thoughts for theoreticians

- ▶ How do you future-proof new constructions?
- ▶ How do you protect against implementation mistakes?
- ▶ How do you defend against politics?