The Next Forty Years of Public-Key Cryptography

Bart Preneel

COSIC KU Leuven and iMinds, Belgium
Bart.Preneel(at)esat.kuleuven.be
September 2014

Thanks to Christiane Peters, Alan Szepieniec, and Frederik Vercauteren

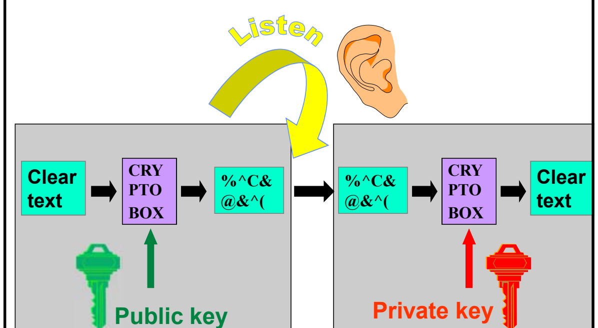
Outline

- Public-key cryptography today
- Attacks on public-key cryptography
- The future: post-quantum crypto
- The future: more than the basics

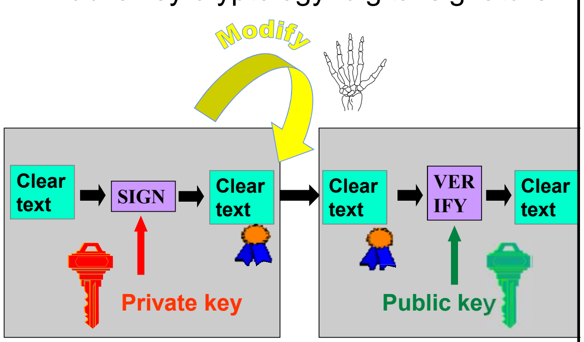
Diffie-Hellman'75
Merkle'75

- Can two people who have never met have a private conversation?
- Is it possible to digitally sign documents?

Public key cryptology: encryption



Public key cryptology: digital signature



Diffie-Hellman public-key agreement

- Before: Alice and Bob have never met and share no secrets; they know public system parameters: a group and a generator α

generate x *compute* α^x	α^x →	*generate* y *compute* α^y
	← α^y	
compute $k=(\alpha^y)^x$		*compute* $k=(\alpha^x)^y$

- After: Alice and Bob share a short term key k unknown to Eve
  - for several groups it is believed to be hard to derive x from α^x (discrete logarithm or DLOG problem)
  - security: Computational DH assumption

RSA ('78)

choose 2 "large" primes p and q
modulus n = p.q
compute $\lambda(n) = \text{lcm}(p-1,q-1)$
choose e relatively prime w.r.t. $\lambda(n)$
compute $d = e^{-1} \bmod \lambda(n)$

public key = (e,n)
private key = d of (p,q)

The security of RSA is based on the "fact" that it is easy to generate two large primes, but that it is hard to factor their product

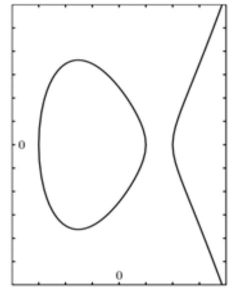
encryption: $c = m^e \bmod n$
decryption: $m = c^d \bmod n$

7

Elliptic Curve Cryptology

Example:
the elliptic curve
$y^2 = x^3-7x+6$
over the field of real numbers R

Advantage: shorter key lengths



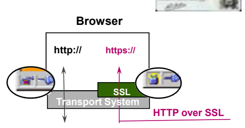
8

Deployment: public-key/hybrid

- PCs/mobile phones/tables (> 3B): automatic updates
- EMV: RSA smart cards (>1B)
  - upgrading to ECC: 2015-2030

- Electronic ID cards and E-passports (~100M)

- TLS/SSL web servers (~10M)
- DNSSEC
- Skype (~500M)
- Bitcoin (~1M)
- The Internet of Things in 2020 (~ 20-50B)



9

Deployment: symmetric key (only)

- GSM/3G/4G – no end to end encryption
- hard disk encryption (BitLocker)
- some payment systems (e.g. Maestro)
- pay TV
- content protection – DRM (DVD, BluRay, iTunes….)

Limited fraction(a few %) of traffic is protected.
A very small fraction of traffic is protected end-to-end with a high security level

10

Deployment of cryptography

- mostly for data and entity authentication
- confidentiality
  - government/military secrets
  - DRM/content protection
  - telco: not end-to-end or with a backdoor
  - hard disk encryption: backdoor
  - most data in the cloud is not encrypted

COMSEC need authenticated encryption/secure channels
  - reordering, replay, deletion of packets
  - protection of **meta-data**

Cryptography is **NOT** used to protect Alice and Bob but to protect the (intellectual) property of corporations

11

All widely used public-key systems rely on 3 problems from algebraic number theory

Integer factorization: RSA (n = p.q)
Discrete **LOG**arithm : Diffie-Hellman, DSA: $y = \alpha^x$
Elliptic Curve **D**iscrete **LOG**arithm, ECDSA: Q = x.P

RSA-1024 ~ DLOG-1024 ~ ECC-146
RSA-2048 ~ DLOG-2048 ~ ECC-206
RSA-4096 ~ DLOG-4096 ~ ECC-282

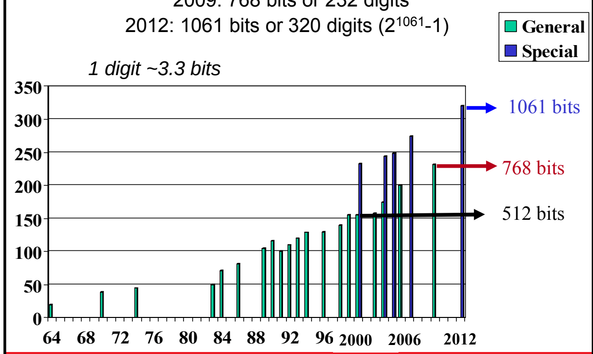
Are these problems hard?

A hard problem is a problem that nobody works on (James L. Massey)

12

Factorisation records (RSA)

2009: 768 bits or 232 digits
2012: 1061 bits or 320 digits ($2^{1061}-1$)



13

NSA and cryptanalysis

Can NSA break
- RSA-512: easily
- RSA-768: definitely
- RSA-1024: likely
- RSA-1536: perhaps
- RSA-2048: who knows

14

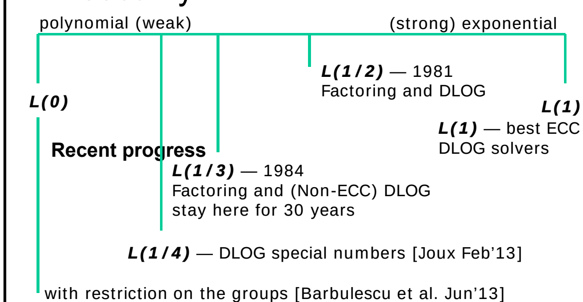
The Cryptocalypse?

2013 breakthrough for DLOG in group of special form

15

Public key crypto security

$L(a)=exp((log_2 n)^a (log_2 log_2 n)^{1-a})$



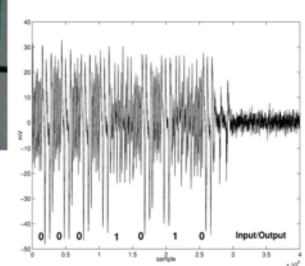
Special form DLOG record: 9234 bits [Granger+'13]
Supersingular binary curves 59-bit security << 128 [Granger+'13]

16

Security in Practice

17

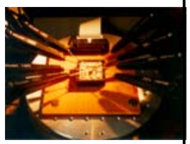
Physics trumps Mathematics



18

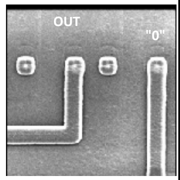
Invasive attacks

Passive: micro-probing



Active: modify circuits

- connect or disconnect security mechanism
  - disconnect security sensors
  - RNG stuck at a fixed value
  - reconstruct blown fuses
- cut or paste tracks with laser or focused ion beam



RNG [www.fa-mal.com]

19

RSA with Chinese Remainder Theorem

[Boneh-DeMillo-Lipton'96]

$s = m^d \bmod pq$
$d1 = d \bmod (p-1)$
$d2 = d \bmod (q-1)$

$s1 = m^{d1} \bmod p$	$s1' \neq m^{d1} \bmod p$
$s2 = m^{d2} \bmod q$	$s2 = m^{d2} \bmod q$
$s = a1\ s1 + a2\ s2 \bmod n$	$s' = a1\ s1' + a2\ s2 \bmod n$

now $\gcd(s-s',n) = q$
since $s = s' \bmod q$ and $s \neq s' \bmod p$

20

Implementation attacks (CHES conference)

Academic

- ever more sophisticated attacks
- broad range of countermeasures: well understood
- new constructions with security proofs: leakage resilience
- cost in practice: 2-100 times more

Industry

- needs security at cost 20-50% more
- return to security by obscurity
- expensive (but confidential) validation program under Common Criteria

21

Many other ways to get the keys
(in particular if you are the NSA)

- Ask for private keys with a security letter
- Substitute public keys
- Put a backdoor in a random number generator that allows to predict outputs

seed → **Pseudo-random number generator (PRNG)** →

Random challenges

22

Dual_EC_DRBG or Dual Elliptic Curve Deterministic Random Bit Generator

- 1 of the 4 PRNGs in NIST SP 800-90A
- Published 2006 based on earlier work by ANSI
- Many warnings about security
  - security proof; but weak if one fails to choose P and Q at random, e.g. Q = d.P for a known d [Brown'06]
  - backdoor [Ferguson-Shumov'07]

Appendix: The security of Dual_EC_DRBG requires that the points P and Q be properly generated. To avoid using potentially weak points, the points specified in Appendix A.1 should be used.

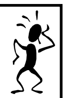
23

Dual_EC_DRBG or Dual Elliptic Curve Deterministic Random Bit Generator

- NSA **Bullrun program:** NSA has been actively working to "Insert vulnerabilities into commercial encryption systems, IT systems, networks, and endpoint communications devices used by targets."
- 10 Sept. 2013, NYT: "Dual EC DRBG standard contains a **backdoor** for the NSA."
- Sept. 2013: NIST "strongly recommends" against the use of dual_EC_DRBG

24

If a large quantum computer can be built...



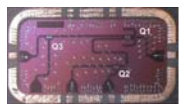
all schemes based on factoring (RSA) and DLOG are insecure [Shor'94]

- including elliptic curve cryptography

symmetric key sizes: x2 [Grover]

hash sizes: unchanged (for collisions)

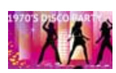
News in Jan. 2014: NSA has spent 85 M$ on research to build a quantum computer



Outline

- Public-key cryptography today
- Attacks on public-key cryptography
- The future: post-quantum crypto
- The future: more than the basics

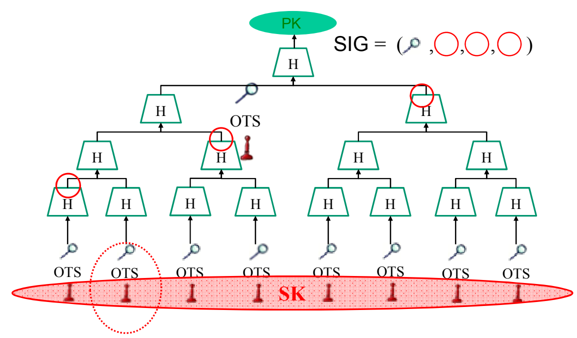
Post-Quantum Cryptography



- Go back to the 1970s
  - digital signatures based on one-way functions
  - public-key encryption based on Error Correcting Coding [McEliece'78]
  - public key encryption based on lattices (inspired by knapsack problems)

- Go back to the 1980s:
  - multivariate polynomial equations
- So far no good quantum algorithms known to break these systems

Hash-Based Signatures



Slide credit: Andreas Hülsing

Hash-Based Signatures: variant XMSS

C Implementation using OpenSSL on Intel(R) Core(TM) i5-2520M CPU @ 2.50GHz with Intel AES-NI [BDH'11]

	Sign (ms)	Verify (ms)	Signature (bit)	Public Key (bit)	Secret Key (byte)	Bit Security	Comment
XMSS-SHA-2	35.60	1.98	16,672	13,600	3,364	157	h = 20, w = 64,
XMSS-AES-NI	0.52	0.07	19,616	7,328	1,684	84	h = 20, w = 4
RSA 2048	3.08	0.09	≤ 2,048	≤ 4,096	≤ 512	87	

Slide credit: Andreas Hülsing

McEliece (1978): code-based public-key crypto

Public key
a random-looking binary linear code given by a matrix *H* weight *w*

Private key
random-looking code is a disguised Goppa code with error-correction capability *w*

Encryption
encode a plaintext as weight-*w* word *e* and send syndrome *s*=*H*·*e*

Decryption
after conversion use standard Goppa-code decoders to determine low-weight solution *e*

Slide credit: Christiane Peters

McEliece security notions

Private key security
Relies on the difficulty of retrieving inner code from public matrix *H* and thus getting access to efficient decoding

Message security
decryption security relies on NP-hardness of the syndrome-decoding problem foro random code - assuming that structure of *H* does not leak
(best known algorithms take exponential time)

Slide credit: Christiane Peters 31

Performance McEliece

C Implementation on Intel Core i5-3210M, Ivy Bridge (encryption times are estimates)

	Decrypt (cycles)	Encrypt (cycles)	Public Key	Secret Key	Bit Security	Comment
RSA-1024	1,340,040	(92,000)	1024 bits	1024 bits	80	
DH binary ECC	77,468	(78,000)	508 bits	508 bits	127	
McEliece	60,493	(73,000)	187 kByte	187 kByte	128	(n,w)= (212,41)

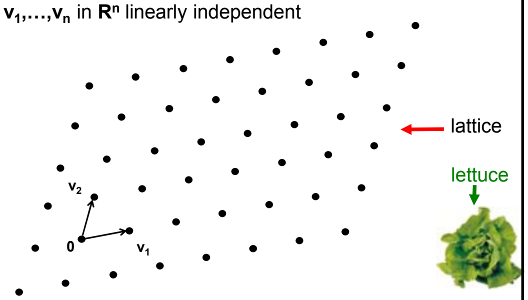
32

Lattices

A **lattice** is a set of points

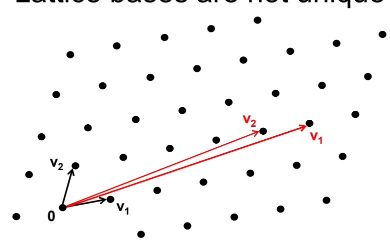
L={a_1v_1+…+a_nv_n | a_i integers}

with **v_1,…,v_n** in **R^n** linearly independent



Slide credit Frederik Vercauteren 33

Lattice bases are not unique

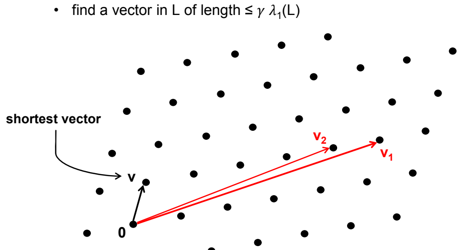


- good basis: short and nearly orthogonal (private key)
- bad basis: long with acute angles (public key)
- fundamental domain = spanned by bases vector

Slide credit Frederik Vercauteren 34

Shortest Vector Problem (SVP)

- SVP: find shortest non-zero vector **v** in lattice L
- Length λ_1(L) of **v** is called the (first) lattice minimum
- SVPγ: relaxed version of SVP
  - find a vector in L of length ≤ $\gamma\ \lambda_1$(L)



Slide credit Frederik Vercauteren 35

Hardness of lattice problems

- SVPγ
  - NP-hard for small $\gamma \sim n^{c/\log\log n}$
- Exact algorithms run in exponential time 2^n
  - [Ajtai-Kumar-Sivakumar'02] [Micciancio-Voulgaris'10]
- Polytime only algorithms for exponential approximations
  - $\gamma \sim 2^{n\ \log\log n/\log n}$
  - ([LLL'82], [Schnorr'87], [Ajtai-Kumar-Sivakumar'02]
- No better quantum algorithms known!

Slide credit Frederik Vercauteren 36

Learning With Errors (LWE)

- $\mathbf{Z}_q^n$ = n-dimensional vectors modulo q, error rate $\alpha << 1$
- Given m vectors $\mathbf{a}_1,\ldots,\mathbf{a}_m$ in $\mathbf{Z}_q^n$
- **Search:** find secret vector **s** in $\mathbf{Z}_q^n$ given "noisy" inner products

$$b_1 = < a_1 , s > + e_1$$
$$b_2 = < a_2 , s > + e_2$$
$$....$$
$$b_m = < a_m , s > + e_m$$

- Errors e_i are taken from Gaussian over **Z** with **deviation** αq

- **Search LWE =** noisy linear algebra modulo q
- m x n matrix **A** with rows $\mathbf{a}_i$: $\mathbf{A}\ \mathbf{s}^t = \mathbf{b}^t + \mathbf{e}^t$

Slide credit Frederik Vercauteren 37

Learning With Errors (LWE)

- $\mathbf{Z}_q^n$ = n-dimensional vectors modulo q, error rate $\alpha << 1$
- Given m vectors $\mathbf{a}_1,\ldots,\mathbf{a}_m$ in $\mathbf{Z}_q^n$
- m x n matrix **A** with rows $\mathbf{a}_i$
- **Decision:** distinguish two distributions

  (A, $\mathbf{b}^t = \mathbf{A}\ \mathbf{s}^t + \mathbf{e}^t$) from uniform distribution (A, $\mathbf{b}^t$)

  - algorithm for decision problem implies algo for search version
  - the secret vector **s** can have entries from the error distribution

- LWE corresponds to BDD on

  L = { **z** in Z^m | $\mathbf{z}^t = \mathbf{A}\ \mathbf{s}^t$ mod q, for some **s** in $\mathbf{Z}_q^n$ }

Slide credit Frederik Vercauteren 38

LWE-based Encryption

- System wide n x n matrix **A** with entries in $\mathbf{Z}_q$
- **Public key:** LWE sample

  (A, $\mathbf{b}^t = \mathbf{A}\ \mathbf{s}^t + \mathbf{e}^t$)

- **Private key:** small LWE secret **s** from error distribution

- **Encryption:** m in {0, 1}
  - generate two small vectors **r, x** with entries from noise distribution
  - ciphertext: **C = (r A + x, <r, b> +** x' **+** m q/2)

- **Decryption:** given ciphertext **C = (c, d)**
  - given **s**, compute **<c, s>** - d ~ m q/2 + small error
  - can easily recover m

Slide credit Frederik Vercauteren 39

LWE-based Encryption: Parameters

- estimate using Bounded Distance Decoding [Liu-Nguyen'13]
- 128-bit security (2^{128} basic ops):
  - dimension n = 256
  - prime q = 7681
  - parameter of Gaussian error distribution ~ 11 (st. dev. $11/\sqrt{2\pi}$)
- public key: 104 Kbyte
- ciphertext: 416 byte

- public key and ciphertext expansion can be reduced with ring version of LWE (structured A instead of random A)
  - hardness related to problems in "ideal" lattices

Slide credit Frederik Vercauteren 40

Key Aspects of Lattice-based Systems

Pros

- efficient and parallizable
  - matrix-vector arithmetic, Fast-Fourier Transform for polynomial multiplication
- worst-case to average-case reductions

Cons

- difficult to find good sampling methods
- difficult to assess exact security
- large keys

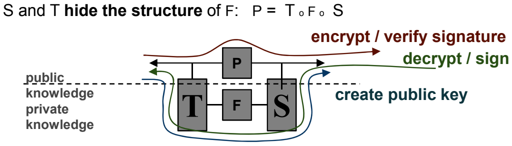
Slide credit: Christiane Peters 41

Multivariate Quadratic Equations

Public Key:
- system of quadratic polynomials P : $F_q^{\,n} \rightarrow F_q^{\,m}$

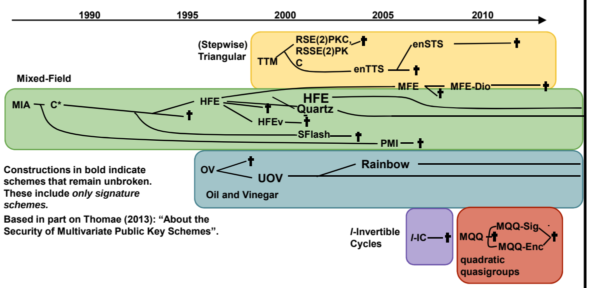
Private Key:
- **affine** transformations T : $F_q^{\,m} \rightarrow F_q^{\,m}$ (on output variables) and S : $F_q^{\,n} \rightarrow F_q^{\,n}$ (on input variables)
- central system of **quadratic** polynomials F : $F_q^{\,n} \rightarrow F_q^{\,m}$ (easily invertible)

S and T **hide the structure** of F: P = T ∘ F ∘ S



Slide credit: Alan Szepeniec 42

Multivariate Quadratic Equations



Slide credit: Alan Szepieniec

43

COMSEC - Communication Security

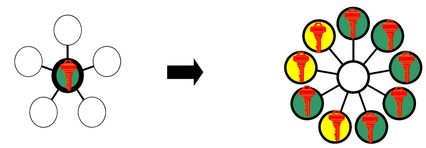
Undermining of end systems (cf. Snowden)

Do **not** move problems to the authenticity of a single public key

Do **not** move problems to a single secret key

– solution: threshold cryptography; proactive cryptography

Do protect meta-data



44

COMPUSEC - Computer Security

Protecting data at rest

– well established solutions for local encryption: Bitlocker, Truecrypt
– infrequently used in cloud
– Achilles heel is key management

45

COMPUSEC - Computer Security

Complex ecosystem developed over 40 years by thousands of people that has many weaknesses

- **Errors** at all levels leading to attacks (think )
  – governments have privileged access to those weaknesses
- Continuous remote **update** needed
  – entity that controls updates is in charge
- Current **defense technologies** (firewall, anti-virus) not very strong
  – cannot resist a motivated attacker
- Not designed to resist **human factor** attacks: coercion, bribery, blackmail
- **Supply chain** of software and hardware vulnerable and hard to defend
  - **backdoors** are hard to detect

46

COMPUSEC - Computer Security

- Simplify to reduce attack surface
- Secure local computation
  - with threshold security
  - Multi Party Computation
  - hardware support: TPM, SMART, Sancus, SGX,…
- Secure and open implementations
- Community driven open audit

47

Reconsider every stage

Crypto design	Kleptography
Hardware/software design	Hardware backdoors
Hardware production	
Firmware/sw impl.	Software backdoors
Device assembly	Adding/modifying
Device shipping	hardware backdoors
Device configuration	Configuration errors
Device update	Backdoor insertion

48

Predictions on the Next 40 Years of Public-Key Cryptography

- **????????????:** Computers, communications, storage are all quantum and all classical cryptography disappears
- **Highly unlikely:** public-key cryptography will disappear completely
  - everything online: symmetric cryptography could make a comeback for many applications (e.g. EMV, web security)
- **Probable**: within 10-20 years massive deployment of post-quantum cryptography (hash-based signatures and lattice-based encryption)
- **Probable:** much more sophisticated protocols with distributed crypto and multi-party computation are more widely used
- **Perhaps:** RSA/DLOG/ECC stays around but with much larger key lengths

Long term security problem is restricted to confidentiality – one can always re-sign if compromise is suspected