

Cryptography, Freedom, and Democracy

How Basic Science Affects Everyone

Nelson H. F. Beebe

Research Professor
University of Utah
Department of Mathematics, 110 LCB
155 S 1400 E RM 233
Salt Lake City, UT 84112-0090
USA

Email: beebe@math.utah.edu, beebe@acm.org,
beebe@computer.org (Internet)
WWW URL: <http://www.math.utah.edu/~beebe>
Telephone: +1 801 581 5254
FAX: +1 801 581 4148

30 November 2005

Near the end of his life, one of the Twentieth Century's most eminent mathematicians wrote in his memoir:

There is one comforting conclusion which is easy for a real mathematician. Real mathematics has no effects on war. No one has yet discovered any warlike purpose to be served by the theory of numbers or relativity, and it seems very unlikely that anyone will do so for many years.

G. H. Hardy, *A Mathematician's Apology*, p. 140 (1940)

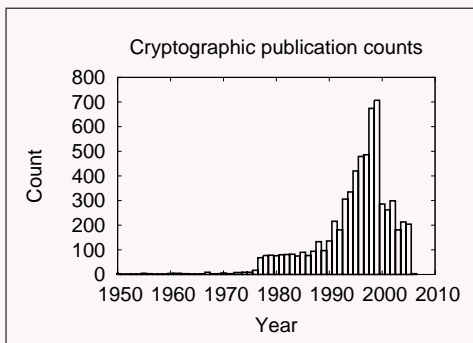
He was wrong!

- Albert Einstein's Special Theory of Relativity (1905), with its famous equation, $E = mc^2$, relates energy, mass, and the speed of light ($c = 299\,792\,458$ m/s (*exact!*) $\approx 186\,282$ miles/s in vacuum).
- Just two years after getting his doctorate, Niels Bohr in Copenhagen, Denmark developed an early quantum theory of the atom in 1913.
- Erwin Schrödinger in Germany discovered the quantum-mechanical wave equation in 1926.
- Otto Hahn and Fritz Strassman in Germany first split the uranium atom by neutron bombardment in 1938. This was confirmed by Lise Meitner and Otto Frisch (Meitner's nephew) in Sweden on December 24, 1938.

- Nazi Germany invaded Poland on September 1, 1939, beginning World War II.
- Japan attacked Pearl Harbor, HI on December 7, 1941, bringing the until-then-neutral USA into the war.
- Manhattan Project began in 1942 at Columbia University Physics Department in New York City (December 6, 1941 by FDR!)
- Manhattan Project later involved University of Chicago, and secret new towns of Hanford, WA, Los Alamos, NM, and Oak Ridge, TN.
- First atomic bomb exploded at the Trinity Site in Alamogordo, NM at 5:29:45 am on July 16, 1945.
- Flight crews trained at Wendover, UT flew from Tinian Island in the Pacific to Japan to drop the Little Boy bomb on August 6, 1945 on Hiroshima, and the Fat Man bomb on Nagasaki on August 9, 1945.
- Japan surrendered August 14, 1945 (formally on September 2), ending World War II.
- Nuclear arms race and the Cold War began shortly thereafter.

Number theory

Whitfield Diffie and Martin Hellman at Stanford University in 1976, and Ralph Merkle at the University of California, Berkeley in 1975 (but unpublished until 1978), independently discovered **public-key cryptography**. Their work was based on some fundamental problems of number theory, and unleashed a flurry of research:



This lecture will discuss why this work matters to every citizen.

- In September 2005, a paper appeared in the *Journal of Cryptology* on **relativistic cryptography**, and a Web search with <http://www.google.com/> found 17 documents with that phrase, the oldest being from 1998. One has the title *Remarks on Mistrustful Quantum and Relativistic Cryptography*, connecting the three basic fields in the introduction to this talk.
- Corrections from both Special Relativity (1905) and General Relativity (1916) are essential for the Global Positioning System on which modern air traffic now depends.

Preliminaries: Some dictionary definitions

code A system of symbols, letters, or words given certain arbitrary meanings, used for transmitting messages requiring secrecy or brevity.

cipher A message written in a secret code.

cryptogram A piece of writing in code or cipher.

cryptography The science of analyzing and deciphering codes and ciphers and cryptograms.

cryptanalysis The analysis and deciphering of cryptographic writings or systems.

encryption To put into code or cipher.

decryption To decode or decipher.

plaintext The unencrypted form of an encrypted message.

ciphertext A text in encrypted form, as opposed to the plain text.

Preliminaries: Some dictionary definitions . . .

prime number A positive whole number not divisible without a remainder by any positive whole number other than itself and one.

For example, the primes up to 100 are:

2 3 5 7 11 13 17 19 23 29 31 37 41

43 47 53 59 61 67 71 73 79 83 89 97

steganography Hiding a secret message within a larger object in such a way that others can not discern the presence or contents of the hidden message.

For example, a message might be hidden within an image by changing the least significant bits to be the message bits.

A cartoonist's view of prime numbers



Simple cryptography: substitution ciphers

Change each letter into another unique letter.

A	B	C	D	E	F	G	H	I	J	K	L	M
↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓
Q	U	Z	M	X	L	K	T	G	P	R	H	O
N	O	P	Q	R	S	T	U	V	W	X	Y	Z
↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓
V	Y	D	E	W	J	S	A	N	C	F	I	B

For example, to encrypt a message, use the rules in that table like this:

plaintext	ATTACK	AT	DAWN
substitute	↓	↓	↓
ciphertext	QSSQZR	QS	MQCV

To decrypt, just reverse the substitution direction:

ciphertext	QSSQZR	QS	MQCV
substitute	↓	↓	↓
plaintext	ATTACK	AT	DAWN

Simple cryptography: substitution ciphers ...

One of the earliest substitution ciphers is the **Caesar cipher** (ca. 50BCE). The substitutions are not to randomly-ordered letters, but rather to the same alphabet shifted circularly by three places.

A	B	C	D	E	F	G	H	I	J	K	L	M
↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓
D	E	F	G	H	I	J	K	L	M	N	O	P
N	O	P	Q	R	S	T	U	V	W	X	Y	Z
↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓
Q	R	S	T	U	V	W	X	Y	Z	A	B	C

Encryption proceeds as before:

plaintext	ATTACK	AT	DAWN
substitute	↓	↓	↓
ciphertext	DWWDFN	DW	GDZQ

Decryption is just the reverse: change ↓ to ↑.

There are two important features of substitution ciphers:

- A **secret key** controls the encryption, either the substitution table (for example, **QUZMXLKTGPRHOVYDEWJSANCFIB**), or for the simpler Caesar cipher, just the number **3** that determines the table shift distance.
- Encryption and decryption are **symmetric**: the same key is used for both. Most cryptographic methods share this property (but *public-key cryptography* does not).

Kerchhoffs' principles of cryptography (1883)

- 1 The system must be practically, if not mathematically, indecipherable.
- 2 It must not be required to be secret, and it must be able to fall into the hands of the enemy without inconvenience (**Kerchhoffs' law**).
- 3 Its key must be communicable and retainable without the help of written notes and changeable or modifiable at the will of the correspondents.
- 4 It must be compatible with the means of communication (most security mechanisms result in message expansion and transform text into nontextual data).
- 5 It must be portable, and its usage and function must not require the concurrence of several people (consider what happens if you log onto a banking site from computer B when your keys are stored on computer A).
- 6 Given the circumstances that command its application, the system must be easy to use, requiring neither mental strain nor the knowledge of a long series of rules to observe.

We have to assume that an attacker has captured our ciphertext. Encryption security then depends primarily on:

- the secrecy of the plaintext,
- the secrecy of the key,
- the complexity of the key (simple keys can be guessed by automated dictionary attacks),
- the quality and strength of the encryption method, and
- the difficulty of cracking captured ciphertext by cryptanalysis.

Security can *sometimes* be improved by:

- hiding word boundaries (concealing **yes** from **no**),
- padding messages to a fixed length with (usually) random characters,
- adding random prefixes and suffixes to messages,
- using multiple levels and/or methods of encryption, and
- changing the key at suitable intervals (daily, hourly, or even with each message).

Frequency analysis

Expected letter frequencies of natural-language text is important for cryptanalysis. Large bodies of English text suggest the order

e t a o i n s h r d l u:

Alice in Wonderland		Hamlet		Roget's Thesaurus		Treasure Island	
19.75%	space	15.70%	space	16.00%	space	18.61%	space
9.40%	e	9.04%	e	8.41%	e	9.28%	e
7.43%	t	7.11%	t	5.81%	a	6.96%	t
6.00%	a	6.53%	o	5.63%	t	6.54%	a
5.69%	o	5.87%	a	5.49%	i	6.03%	o
5.22%	i	5.09%	i	5.34%	n	5.31%	n
4.92%	h	4.95%	s	5.27%	o	4.95%	h
4.84%	n	4.92%	h	4.87%	r	4.95%	i
4.46%	s	4.90%	n	4.36%	s	4.67%	s
3.86%	r	4.63%	r	3.84%	,	4.26%	r
3.36%	d	3.71%	l	3.41%	c	3.77%	d
3.24%	l	3.06%	d	3.33%	l	3.19%	l
2.40%	u	2.70%	u	2.65%	u	2.28%	u

Comments on cryptanalysis

- The more ciphertext that is available, the easier it is for substitution ciphers to be cracked by letter-frequency analysis.
- Serious cryptanalysts have automated tools for tackling *all* of the common and historical encryption schemes, and most can be broken very quickly.
- Certain government agencies devote very large resources to cryptanalysis. The US National Security Agency is believed to be the world's largest employer of mathematicians, have some of the world's largest computer systems, and its budget is secret.
- Successful cracks are not announced by such agencies, nor is their research published.
- US and Britain monitor and analyze all transatlantic telephone and network traffic.

Weaknesses of simple cryptographic methods

- With enough ciphertext, all of the standard ones are easily crackable.
- Unless the key is changed, a given plaintext letter is always converted to the same ciphertext letter, facilitating frequency-analysis attacks.
- Key reuse makes cryptanalysis easier.
- Chosen plaintext makes cryptanalysis easier (e.g., encryption of disinformation).
- Keys may be captured at either end of the communications channel, without the other end detecting the capture, compromising all future traffic.
- Keys must be shared by sender and receiver: this is the **most serious drawback** (e.g., key changes for army headquarters and troops in the battlefield, or naval command and a submarine).
- Eavesdropping is rarely detectable (photons and electrons cannot be tagged).
- Traffic analysis can still reveal important information, even if the traffic itself cannot be understood by the attacker.

Stream and block ciphers

- Simple encryption methods work on a character (or bit) at a time; they are **stream ciphers**.
- In stream ciphers, a particular character is encrypted identically, no matter where it appears in the data stream.
- A transmission error in a single character affects only that character.
- Better methods work on groups of characters (or bits) at a time; they are **block ciphers**.
- In a block cipher, the encryption of a particular character depends on all others in the same block.
- Thus, in a block method, a particular character will usually be encrypted differently, depending on its surroundings.
- A transmission error in a single character affects the entire block.
- Block methods therefore require reliable communications.
- The best modern encryption methods are usually block ciphers.

Uncrackable encryption method: the one-time pad

Cryptanalysis is possible whenever there are patterns in the encryption of plaintext to ciphertext. The only way to prevent cryptanalysis is to use a **different** encryption for each plaintext letter, since that destroys all patterns.

A **one-time pad** satisfies this requirement. For example, use successive letters of text from a mutually-agreed-on book (the **key**) to determine the shift count of a Caesar-like substitution cipher:

Call me Ishmael. Some years ago—never mind how long precisely—having little or no money in my purse, and nothing particular to interest me on shore, I thought I would sail about a little and see the watery part of the world.

Herman Melville, *Moby Dick*, London (1851)

Weaknesses of our one-time pad

- Unfortunately, when a book of natural-language text provides the one-time pad, there are still patterns present that can allow cryptanalysis (e.g., **and**, **I**, **little**, **me**, and **the** occur twice, and some words have repeated letters (**ee**, **ll**, and **tt**)).
- What does one do when the book is exhausted? The pad cannot safely be reused.
- What is needed is a **completely-random string of letters** of **unlimited length** for the one-time pad.
- A computer method for generating random numbers requires a starting number, called the **seed**, that serves as the **encryption key**.

Example of the one-time pad

The encryption does not reveal message length, although it **does** reveal common plaintext prefixes:

`encrypt(123, "A")`

```
2b04aa0f ef15ce59 654a0dc6 ba409618 daef6924 5729580b  
af3af319 f579b0bc
```

`encrypt(123, "AB")`

```
2b47315b 22fdc9f1 b90d4fdb 1eb8302a 4944eddb e7dd1bff  
8d0d1f10 1e46b93c
```

`encrypt(123, "ABC")`

```
2b47752c 286a4724 40bf188f c08caffa 1007d4cc 2c2495f9  
cd999566 abfe0c2d
```

`encrypt(123, "ABCD")`

```
2b477571 f970b4a2 7346ca58 742e8379 e0ce97b3 1d69dc73  
c7d921dc 018bc480
```

Example of the one-time pad ...

The encryption does not reveal letter repetitions:

```
encrypt(123, "AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA")
2b46736e 3b83cd28 777d88c8 ad1b12dc c28010ef 407d3513
e1ed75bc 5737fd71 6e68fb7d 4ac31248 94f21f9f d009455f
6d299f
```

Now encrypt a famous message from American revolutionary history:

```
ciphertext = encrypt(123, \
"One if by land, two if by sea: Paul Revere's Ride, 16 April 1775")
println ciphertext
3973974d 63a8ac49 af5cb3e8 da3efdbb f5b63ece 68a21434
19cca7e0 7730dc80 8e9c265c 5be7476c c51605d1 af1a6d82
9114c057 620da15b 0670bb1d 3c95c30b ed
```


Example of the one-time pad ...

Attempt to decrypt the ciphertext with a nearby key. Decryption **does** reveal the message length, although that flaw could easily be fixed:

```
decrypt(122, ciphertext)
```

```
?^?/?)?D?fN&???w??V???Gj5?????(????1???J???i?i)y?I?-G?????b?o???X?
```

Attempt to decrypt the ciphertext with the correct key:

```
decrypt(123, ciphertext)
```

```
One if by land, two if by sea: Paul Revere's Ride, 16 April 1775
```

Attempt to decrypt the ciphertext with another nearby key:

```
decrypt(124, ciphertext)
```

```
??$???W?????N?????????!?Z?U?????????Q?????????3?B}'<?0 ?P5%??VdNv??kS??
```

Lesson: a nearby key is as useless as a faraway key: almost-right isn't good enough.

Limitations of the one-time pad

- Although very good methods are now known for generating random numbers on a computer, they are always produced by a specific recipe that introduces patterns that can aid cryptanalysis, and most of the popular methods have been cracked.
- Truly-unpredictable sources of random numbers, such as radioactive decay, cannot be replicated for the sender and receiver.
- The problem of secure key distribution remains.

Public-key cryptography

- Public-key cryptography solves the key-distribution problem. Instead of a **single shared secret key**, each participant has a pair of keys: a **public key**, and a companion **private key**.
- The keys are related, such as two very large prime numbers whose product is hard to factorize. However, that is mathematically a very hard problem: given one of the keys, it is **computationally infeasible** to determine the other.
- Such problems are sometimes called **one way trap doors**.
- Easy to put needle in haystack, but much harder to remove it.

Public-key cryptography and prime numbers

- Prime factorization of small numbers is easy:

$$\begin{aligned}99 &= 3 \times 3 \times 11 \\6860 &= 2 \times 2 \times 5 \times 7 \times 7 \times 7 \\62271 &= 3 \times 3 \times 11 \times 17 \times 37 \\62273 &= 62273 \quad \text{prime number} \\97272 &= 2 \times 2 \times 2 \times 3 \times 3 \times 7 \times 193\end{aligned}$$

- Prime factorization of big numbers is hard:

$$\begin{aligned}1447473570262981491527798 &= 2 \times 109 \times 3687427 \times 12523837 \times 143778289 \\8992987500442157627511191 &= 19 \times 318023201 \times 1488303778195789 \\8992987500442157627511193 &= 8992987500442157627511193 \quad \text{prime number} \\17054727660401396805027270 &= 2082815984930 \times 8188302655539\end{aligned}$$

- Brute-force factorization of an N -digit number could require trying all factors up to size $N/2$ digits.

Using public key cryptography

Alice and Bob communicate securely as follows:

- Alice encrypts her plaintext with Bob's public key and sends it to him.
- Bob decrypts Alice's ciphertext with his private key to recover her plaintext.
- Bob encrypts his response with Alice's public key and sends it to her.
- Alice decrypts Bob's ciphertext with her private key to recover his plaintext.

Security and uses of public-key cryptography

- An attacker cannot claim to be either Alice or Bob, since their private keys are not known to him.
- The technique can easily be extended to allow **unforgeable digital signatures** (if private keys remain secret).
- The mathematics of prime factorization has received a lot of study, and is believed to be intractable for large numbers (200+ digits).
- Other public-key methods based on elliptic curves or discrete logarithms provide a failover should a mathematical breakthrough uncover a fast way to crack factorization-based methods.
- Public-key methods are not practical for routine high-speed communication. They are therefore used to communicate long randomly-chosen keys for faster symmetric methods that are believed to be secure, e.g., **NIST Advanced Encryption Standard (AES)**.
- Examples include **secure shell** on Unix systems, **https://...** Web connections, and some recent network protocols.

Cryptography and the citizen

- Public-key cryptography solves the key-exchange problem, and allows use of hard-to-crack random keys for symmetric methods.
- It can allow individuals to communicate with considerable confidence that no one, not even **secret government agencies** with powerful computers, can crack their communications.
- Private keys are still subject to compromise, such as by carelessness, eavesdropping, threat of violence, or force of law.
- Verification of ownership of public keys remains a problem. Why should Alice believe Bob's public key is really his, unless she knows him personally, and got the key directly from him?
- Public-key certification services by large corporations, such as Verisign, are advocated by some, but this just transfers the trust problem to another large organization over which you have no control, and little confidence in.
- Registration of public keys in a number of different key servers scattered around the world makes it harder to forge a public key.

Alas, no.

- Bruce Schneier wrote two editions of a famous book, **Applied Cryptography**, and with Niels Ferguson, co-authored another, **Practical Cryptography**, describing the mathematics and computer science behind cryptography.
- He wrote two more books, **Secrets & Lies**, and **Beyond Fear**, that deal with the social aspects of security and privacy.
- In modern computer systems, plaintext can be recovered by encryption-key compromise, by capturing data before encryption (e.g., keyboard sniffer, screen images, or keyboard sounds), by trapping data after decryption, or by cracking ciphertext encrypted with weak methods (simple passwords, Bluetooth, WEP on wireless networks, Microsoft Windows passwords and protocols, cell phones, ...).

Can cryptography ensure privacy? ...

- Faulty software implementations of cryptographic methods, communications, and protocols have reduced or eliminated security in far too many cases.
- Only well-studied publicly-available encryption techniques believed to be secure by the cryptographic research community are trustworthy.
- Beware of **“security by obscurity”** and **“proprietary encryption techniques”**.
- Absence of a published successful attack against an encryption method **does not mean that it is secure**. Only published reports of repeated failed attacks and mathematical analysis can give confidence in its security.
- Storage of encrypted data is perilous: if you forget the key, or an employee leaves with the key, your data is lost, compromised, or could be held hostage.
- If an attacker learns your encryption key, your traffic or data may be monitored without your knowledge.

Dangers of technology

- Radio-frequency identification (RFID) tags on people (passports, clothing, embedded under skin, . . .) endanger privacy and can even make them targets for theft or violent attacks.
- Triangulation of cell-phone signals can locate and track everyone carrying such a phone (500M manufactured in 2004).
- Biometric identifiers end up in shared databases with lifelong effect (e.g., fingerprinting and photographing of foreign visitors to US, or all newborns).
- Health-care records and genetic history could lead to insurance premium increases or even denial of coverage.
- Airport watch lists of personal names can result in travel blocks on innocent people.
- High-resolution beneath-clothing scanning of travelers at airports.
- Computer-based facial recognition has high rate of false positives.

Dangers of technology . . .

- Video surveillance cameras in public and corporate areas, and even some homes. And now in cell phones.
- Records of **all** telephone and email traffic to be maintained by European Union for possible prosecution.
- Records of library borrowing demanded by law enforcement.
- Web browser caches and history files.
- Highway toll-booth scanners read vehicle ID and bill the owner.
- Photocop on every corner?
- Chips in World Cup soccer balls used to determine goal entry.
- VCR and DVD players that prevent skipping commercials.
- Region coding in DVDs.

Dangers of technology . . .

- British proposal to put cameras with automatic number-plate recognition **every 400 yards on motorways, as well as at supermarkets, petrol stations and in town centres.**
- British number plates encode vehicle year, plates in many US states include county, dealer tag identifies city, and garage oil-change label identifies neighborhood.
- Revision histories and user information embedded inside complex document formats from office application software.
- Cleartext and otherwise unsecured or insecure wireless networks, including cell phones, Bluetooth, garage/car door openers, . . .
- Unicode characters in Internet hostnames (Internationalized Domain Names) can be deceptive.
- Color printer output encoding printer serial number and time stamp.
- Thieves use Bluetooth phones to find Bluetooth-enabled laptops in parked cars, and then steal the laptops.

Reread George Orwell's book *1984*: **Big Brother** is watching you.

Freedom and democracy

- Electronic communication over wired and wireless channels is now the norm in many countries.
- Free communication between individuals and groups enhances freedom and democracy, and makes it difficult for dictatorships to survive.
- Modern cryptography allows individuals, organizations, and governments to keep private data secure against outside eavesdropping.
- Traffic analysis can still reveal critical information.
- Legislation in some countries makes use of cryptography a crime or treats cryptography (both research and software) as a weapon subject to prepublication review or export controls, or requires individuals to surrender encryption keys to law enforcement or to a government escrow agency (e.g., the US Clipper Chip proposals of the 1990s).

- Cryptography and computer systems are advanced technologies that are **poorly understood** by the general public, and the executive, legislative, and judicial branches of governments.
- Because computers are new and glitzy, there is a tendency to accept and adopt computer technology as if it were reliable, when it frequently is not.
- Electronic voting systems are being widely adopted in many countries, yet they are **impossible to secure, trust, or audit**.
- While elections with paper ballots can be subverted in a few precincts, with electronic voting, the secret ballot and entire national elections are at risk.

Conclusions and lessons

- Learn the limitations of technology, and fight attempts to use it inappropriately.
- The social and technological impact of basic science research is generally impossible to predict until years after the research is done.
- Reject electronic voting.
- Guard against government attempts to destroy freedom by imposing controls and monitoring of citizens in the name of security (e.g., US Patriot Act, and unlimited detention without legal counsel or trial).
- Oppose database aggregation, and excessive collection of unnecessary data that violates your privacy and your economic security.

Where to learn more

- Read Bruce Schneier's books, especially **Secrets & Lies**.
- Subscribe to the monthly **crypto-gram** newsletter.
- Visit Web sites like:
 - <http://www.fsf.org/> (Free Software Foundation),
 - <http://www.gnu.org/> (GNU Project),
 - <http://www.cdt.org/> (Center for Democracy and Technology),
 - <http://verifiedvoting.org/>,
 - <http://www.acm.org/serving/> (ACM site on Computing & Public Policy, with link to Risks Forum),
 - <https://www.ieeecomunities.org/securityandprivacy> (IEEE Security & Privacy forum and journal),
 - <http://www.math.utah.edu/~beebe/> (notes on PGP, and slides and handouts for talks), and
 - <http://www.math.utah.edu/pub/tex/bib> (bibliographies on cryptography, including journals: **Cryptologia** covers history of field).