Software and Security Engineering

Lecture 5

Martin Kleppmann

martin.kleppmann@cst.cam.ac.uk With many thanks to Ross Anderson and Alastair Beresford

Symmetric key cryptography requires careful sharing of keys

i	malasulars				Dias	Dalla		Stedierverbindungen											henngruppen							
	Hom I	we	13enia	ç.e	ning	lienn	ng	en i	ber Um	kehrmo	ilse		2	,	4	5	6	7		9	10					
849	31	1	v	111	14	69	24					SZ	OT	DV	KU	FO	МҮ	EW	JN	IX	LQ	wny	dgy	exb	rzg.	
649	30	IV	111	п	05	26	02					15	EV	МΧ	RW	DT	UZ	JQ	A0	СН	NY	kti	acw	251	0 SW	
849	20	111	11	1	12	24	03	KM	XA	PZ	00	DJ	ΤA	CV	10	ER	QS	LW	PZ.	FN	BH	100	acn	0vw	WVG	
849	28	11	111	v	05	90	16	DI	CN	BR	PV	CR	FV	ΑI	DK	OT	MQ	EU	ВХ	LP	GJ	110	cld	ude	rzn	
649	27	111	1	1V	11	03	07	LT	EQ	HS	UW	DY	IN	BV	OR	MA	ΓO	PP	НТ	EX	UW	woj	100	VCL	uis	
840	26	1	11	v	17	22	19					VZ	AL	RT	ко	CO	EI	BJ	DU	FS	HP	xle	gbo	uev	rxm	
649	25	IV	111	· I	08	25	12					; OR	PV	AD	IT	PK	HJ	LZ	NS	EQ	CW	ouc	unq	uew	414	
649	23	v	1	IV	05	18	14					I TY	AS	٥W	ΚV	JM	DR	нх	GL	CZ	NU	k p 1	rwi	ver udf	tio	
649	23	IV	11	I	24	12	04					QV	FR	AK	EO	DH	CJ	MZ	SX	GN	LT	ebn	rwm		www	
649	22	11	1V	v	01	09	21	10	AS	DV	OL	FJ	ES	IM	RX	LV	AY	00	BO	WZ	CN	Jdc	del	mwf	wvf	
849	21	1	v	11	13	05	19	FT	ox	EZ	СН	RU	HL	FY	os	0Z	DM	AW	CE	TV	mw.	1pw	cef	nvo	vsh	
649	20	111	1V	v	24	01	10	MR	KN	BQ	PW	DF	MO	QZ	A0	RY	SV	JL	0 A	15	GI	jqu	fpx	iwg	tig	
649	19	v	111	1	17	25	20					OX	PR	Fn	WI	DL	UM DY	ND.	pc	1.0	BD	152	=bw	vcj	rxn	
649	18	IV	11	v	15	23	26					EJ	- 01	IV	AQ.	A.M.	I A	MI	AP	IP	BH	mae	hzi '	SOR	ysi	
649	17	1	IV	1!	21	10	06					- 18	K.L	D1	MR	BY	¥7	05	PU	FO	CT	tdp	dhb	fkb	uiv	
649	16	V	11	111	08	16	13					HM	10	MB	OW	LY	AJ	BO	co	IP	NT	ldw	hzj	soh	wvg	
649	15	11	IV	I	01	03	07	1.5				. 04	IP	KS	TY	HZ	PL	AX	BT	CQ	NV	imz	noa	tjv	xtk	
649	14	IV	1	v	15	11	05	IA	BT	MV	но	LY	AG	KM	BR	10	JU	HV	SW	ET	CX	zgr	dgz	gjo	ryg 🔫	
649	13	1 1	111	11	13	20	03	FW	EL	DO	KN	1 MO	BP	CY	RZ	KX	AN	JT	DG	IL	FW	zdy	rkf	tjw	xt1	
649	12	. V	15	IV	18	10	0/	RZ	QQ	CP	SX	KN	UY	HR	PW	FM	во	EZ	QT	DX	JV	zea	rjy	soi	wvh	
649	11	1 11	IV	III	0	20	01					LR	IK	MS	QU	HW	PT	00	VX	FZ	EN	1 r c	zbx	vbm	rxo	
649	.10	III	v	10	12	0	0					QY	BS	LN	KT	AP	IU	DW	HO	RV	JZ	edj	eyr	vby	tin	
649	19	- V		· · · W		10	0 25	-			-	FI	NQ	SY	CU	BZ	AH	EL	ТX	DO	KP	yiz	dha	ekc	th	
649	8	- 10	11	U.	0	0	3 22					. UX	12	HN	BK	QQ	CP	FT	JY	MW	AR	lan	dgb	zsj	woi	
846	1	- 1	IV	v	1	1 11	B 14		-			. DQ	GU	BW	NP	HK	AZ	CI	PO	JX	VY	120	cit	ZSX	wai	
041	1.0	- W	1	IV	12	3 0	2 2	IL	AP	EU	но	MV	CL	OK	OQ	BI	FU	HS	PX	NW	EI	1 Ju	car	1 ye	uib	
041		1	**		1.0			QT	WZ	KV.	GN	. AC	BL	OZ	EK	QW	OP	50	DH	JM	TX	150	208	,		

Photo source: https://commons.wikimedia.org/wiki/File:Enigma_keylist_3_rotor.jpg

A list of keys for a German Enigma cipher machine.

English translation of text along the top (from Wikipedia):

"Secret Command Document! Every individual key setting is secret. Forbidden to bring on aircraft.

Luftwaffe Machine Key No.649

Attention! Key material must not fall into enemy hands intact. In case of danger destroy thoroughly and early."

Public key cryptography

Allows two parties with no prior knowledge of each other to jointly establish a shared secret key over an insecure channel

Examples include Diffie-Hellman and RSA

Diffie Hellman revision

Alice and Bob publicly agree to use p = 23, g = 5

- 1. Alice chooses secret integer a = 4, then A \rightarrow B: g^a mod p = 5⁴ mod 23 = 4
- 2. Bob chooses secret integer b = 3, then B \rightarrow A: g^b mod p = 5³ mod 23 = 10
- 3. Alice computes $10^4 \mod 23 = 18$
- 4. Bob computes $4^3 \mod 23 = 18$

Alice and Bob now agree the secret integer is 18

Example derived from https://en.wikipedia.org/wiki/Diffie-Hellman key exchange

You saw Diffie Hellman in Discrete Maths. This simple version uses a multiplicative group of integers modulo p, where p is prime and g is a primitive root modulo p The values of p and g are chosen in this way to ensure that the resulting shared secret can take on any value from 1 to p-1.

This protocol has a significant limitation: it is susceptible to a person-in-the-middle attack.

Physical public key crypto with locks

- Anthony sends a message in a box to Brutus. Since the messenger is loyal to Caesar, Anthony puts a padlock on it
- Brutus adds his own padlock and sends it back to Anthony
- Anthony removes his padlock and sends it to Brutus, who can now unlock it

Is this secure?

109

Anthony wants to kill Caesar, but needs Brutus' help to do so. How can Anthony send a message to Brutus yet not let the messenger read the message? This proposal is insecure: it is vulnerable to a MITM attack, as is Diffie Hellman.

Here Anthony has shared the secret message with someone, but Anthony doesn't know who it is!



More on this in the Part IB Security course and Part II Cryptography course. Require knowledge at this point is as stated above and expanded on in the lecture.

Note that asymmetric public-key crypto has the same problem as Diffie-Hellman: how do you know that you have the right public key for Alice and you are not subject to a MITM attack?



Once public key crypto is discovered, people then looked for ways to use it. Background: Needham went to California every summer to work at Xerox Parc. He got a preprint of the RSA paper and decided to apply it to the problem on the Xerox network computer project. Kerberos, discussed earlier, was derived from the Needham-Shroeder protocol, and in 1978 Needham proposed the following publickey variant of the protocol.

This version does not require an online server, Sam. Instead the nodes now need the long-term public keys of each other. Here, KA and KB are the *public* keys of A and B respectively, and the aim is to use these in order to derive a symmetric session key between A and B (symmetric cryptography is computationally cheaper).

MITM attack found 18 years later

_А , N _B } _{КА}

The fix is explicitness. Put all names in all messages.

112

Here Charlie can pretend to be Alice when talking to Bob (line 2). Doing so means that Charlie gets N_A (line 1) as well as N_B (line 5). Now Alice correctly believes she's talking to Charlie, but Bob believes he's talking to Alice, when in fact he is talking to Charlie. Bob may therefore subsequently share some information that was intended only for Alice, but Charlie can intercept it.

Don't beat yourself up with if you didn't spot the problem in the protocol. It took 18 years before someone noticed the problem as shown on this slide. (Gavin Lowe: Breaking and fixing the Needham-Schroeder Public-Key Protocol using FDR. TACAS 1996)

Binding keys to principals is hard

- Physically install binding on machines
 - IPSEC, SSH

• Trust on first use; optionally verify later

- SSH, Signal, simple Bluetooth pairing
- Use certificates with trusted certificate authority
 - Sam signs certificate to bind Alice's key with her name
 - Certificate = sig_s{A, K_A, Timestamp, Length}
 - Basis of Transport Layer Security (TLS) as used in HTTPS
- Use certificate pinning inside an app
 - Used by some smartphone apps

Transport Layer Security (TLS)

- Uses public key cryptography and certificates to establish a secure channel between two machines
- Various efforts to prove correctness
 - Core of TLS 1.0 verified by Larry Paulson, 1999
 - TLS 1.3 (2018) designed alongside formal verification
- Nevertheless, many flaws have been found
 - Interactions between protocol extensions, implementation bugs, side-channel (timing) attacks, downgrade to old protocol versions, ...
- Often a large number of root certificate authorities. Are these all trustworthy?

Earlier versions of this protocol were called Secure Sockets Layer (SSL). There's been around one bug every year in TLS since 1999. The first series of attacks were timing attacks: look at how long it takes a server to respond and use this to determine certain bits of the key. The challenge here is that compilers and security engineers fight. Compilers attempt to make code as fast as possible, and may optimise away "make work" inserted by security engineers who are attempting to ensure constanttime execution for critical operations. There are many more technical hacks here, but these are for later courses.

Another major problem is that it's really hard to fix bugs when found. In order to change the protocol you need to make changes to both the client and the server. This is hard for the Web since you have to upgrade both all web browsers and all web servers and no single party is in control of the overall ecosystem. There are poor incentives. There are 187 root certificates installed on my Mac. Web browsers typically trust all of them, and any of these certificates may be used to license other providers with the power to create further certificates for arbitrary domains.

DigiNotar went bust after issuing bogus certificates (2011)

- Dutch certificate authority
- More than 300,000 Iranian Gmail users targeted
- More than 500 fake certificates issued
- Major web browsers blacklisted all DigiNotar certs
- Today: Certificate Transparency to detect incorrectly issued certificates

Iranian Gmail users were found to have been given fake certificates for Gmail, allowing a MITM attack to take place. Further investigation revealed that over 500 fake certificates were issued. No public investigation provides conclusive proof of all steps in the process, but the Iranian Government and the NSA have both been suggested as potential attackers. The behaviour of governments here has a significant influence on the security of everyone else. The cryptowars of the 1990s, where governments attempted to mandate exceptional access to encrypted key material, are being revisited. See: Ableson et al. Keys Under Doormats: mandating insecurity by requiring government access to all

data and communications. https://www.schneier.com/academic/paperfiles/paper-keys-under-doormats-CSAIL.pdf

Further reading: https://en.wikipedia.org/wiki/DigiNotar



Look at the security rating site OpenSSL Labs for the Department's certificate. The landscape here is very complex. You need a detailed tool to check whether your certificate and setup has all the appropriate defences deployed for the various flaws found in the protocol over the years. Note the provision of client compatibility too.



Even if a protocol is secure in isolation, it might fall down if the same key can be reused in multiple protocols. For example, say you have a challenge-response protocol where a service can send a user a random string X and ask them to produce a digital signature over X to authenticate the user. If the same key can be used for both age verification and authorising a purchase, a challenge for a purchase may be replayed in an age verification context. This problem can be prevented by not using the same key for multiple purposes, or by including an indication of the purpose as part of the signed data (e.g. signing the string "age verification response for challenge X" rather than just "X"), so that the signature cannot be reused for another purpose.



Maurice Wilkes: "It suddenly occurred to me when I was at the corner of the stairs, that I would spend a large part of my life discovering bugs in my own programs."

The first documented use of the term "bug" for a technical malfunction was by Thomas Edison; In the year 1878 he mentioned the term in a private letter. This counters an oft-mentioned view that the term bug is derived from a moth getting trapped in a computer, although perhaps this latter event popularised the term. For further information, see <u>https://en.wikipedia.org/wiki/Software_bug</u>

Patriot missile failures in Gulf War I



German Air Force; CC-BY-SA, Darkone, Wikipedia



Afgan National Army; PD, Davric, Wikipedia

- Failed to intercept an Iraqi Scud missile in first Gulf War on 25th February 1991
- Scud struck US barracks in Dhahran; 28 dead
- Other Scuds hit Saudi Arabia, Israel

119

The MIM-104 Patriot is a surface-to-air missile (SAM) system, the primary of its kind used by the United States Army and several allied nations. The picture on the left is a Patriot system used by the German Air Force, August 2005 (<u>https://en.wikipedia.org/wiki/MIM-104_Patriot</u>). The picture on the right is of a Scud missile and launcher in use by the Afgan National Army.

Caused by arithmetic bug

- System measured time in 1/10 sec, truncated from 0.0001100110011..._b
- Accuracy upgraded as system upgraded from airdefence to anti-ballistic-missile defence
- Code not upgraded everywhere (assembly)
- Modules out by 1/3rd sec after 100h operation
- Not found in testing as spec only called for 4h tests Lesson: Critical system failures are typically multifactorial

2020: Boeing 787 must be rebooted every 51 days

As you will know from the Numerical Analysis course, not all decimal fractions are precisely representable as binary floating-point numbers.

System was upgraded from anti-aircraft to anti-ballistic missile. This required an increase in accuracy since ballistic missiles such as the Scud travel much faster than aircraft. Unfortunately the code was not updated everywhere. This meant that different modules (some with upgraded accuracy, some not) then fell out of sync with each other, resulting in the failure of the Patriot system to effectively target Scud missiles. This problem was not caught by static analysis tools since the code was written in assembly, and therefore there was no high-level language features such as a strong type system which could have helped. Testing was also inadequate – missile defence systems are often operated continuously for hundreds of hours, yet the testing regime only called for testing over a 4-hour period. Short-term solution was to reboot Patriot every 4 hours until the underlying cause was determined.



Java supports implicit type conversion or *coercion* from primitive integers to Strings. This is typically helpful, however implicit type conversion interacts with implicit operator precedence in the above example, leading to different outcomes for what initially appear to be quite similar expressions. Removing all implicit type conversion may also result in (different) errors since programmers may then insert explicit type conversions which themselves might be problematic.

Further reading: Joshua Bloch and Neal Gafter, Java Puzzlers: Traps, Pitfalls, and Corner Cases, Addison-Wesley. <u>http://www.javapuzzlers.com/</u>

```
Apple's goto fail bug (2014)
static OSStatus SSLVerifySignedServerKeyExchange(SSLContext *ctx,
                bool isRsa, SSLBuffer signedParams,
                uint8 t *signature, UInt16 signatureLen)
{
   OSStatus err;
   //...
   if ((err = SSLHashSHA1.update(&hashCtx, &serverRandom)) != 0)
      goto fail;
   if ((err = SSLHashSHA1.update(&hashCtx, &signedParams)) != 0)
      goto fail;
      goto fail; //error: this line should not exist
   if ((err = SSLHashSHA1.final(&hashCtx, &hashOut)) != 0)
      goto fail;
   //...
   fail:
      SSLFreeBuffer(&signedHashes);
      SSLFreeBuffer(&hashCtx);
      return err;
                                                           122
}
```

This is a control-flow (logic) bug. Note the two consecutive lines containing "goto fail"; the second is erroneous and the control flow therefore unconditionally executes the code at the "fail" label. It's not clear how this failure was introduced. Perhaps it was an erroneous merge on a commit, either automated or manual by a user. Better unit tests might have helped.

Further reading: https://www.imperialviolet.org/2014/02/22/applebug.html



This is another logic bug. The heartbeat feature allowed either the client or the server to ask the other party to reply with a specified message of a given length after a period of time, allowing the requesting party to know that the other was still online and available.



Unfortunately the requesting party could claim the provided message was much larger than reality. This led to a buffer over-read vulnerability: the requesting party would receive their message appended with any additional contents found in the server or clients memory. The bug's name derives from heartbeat. NB: In the absence of malice, the code worked just fine.

Further reading: https://en.wikipedia.org/wiki/Heartbleed

Heartbleed allows clients to read the contents of server memory

Therefore a malicious client could read:

- Secret keys of any TLS certificates used by server
- User creds such as email address and passwords
- Confidential business documents
- Personal data

The attack left no trace of use in server logs

125

The potential impact of this vulnerability is huge. Potentially the entire contents of the server's process address space were accessible.

One significant risk was for webservers connected to the public Internet. Since the attack left no trace of use in server logs, this meant that all servers needed to not only upgrade their software to fix the vulnerability, but to replace all important key material. TLS certificates in use by the server are an important example, since the private keys may have been compromised. Ideally all user passwords should have been replaced too as these may have been compromised, but the risk of this type of failure depends on details of any implementation.

Notification and clean-up difficult

12 th March 2012	Bug introduced (OpenSSL 1.0.1)
1 st April 2014	Google secretly reports vuln
3 rd April 2014	Codenomicon reports vuln
7 th April 2014	Fix released
7 th April 2014	Public announcement
9 th May 2014	57% of website still using old TLS certificates
20 th May 2014	1.5% of 800,000 most popular websites still vulnerable

The original flaw was introduced into the source code repository for OpenSSL on 31st December 2011, and was released in OpenSSL in version 1.0.1 on 12th March 2012. The bug appears to have been found by multiple people, including members of the security team at Google who produced a fix which appeared on RedHat's issue tracker on 21st March 2014. Codenomicon also discovered the problem independently and reported on 3rd April 2014. [Dates sourced from https://en.wikipedia.org/wiki/Heartbleed]

A significant issue with notification is it was essentially impossible to do so quietly: the number of servers and clients which needed fixing is simply too large. Another problem is that many server operators did not realise that they may have been compromised and therefore did not replace their certificates (potentially allowing a MITM attack on all connections, and in the absence of a version of the protocol with forward secrecy, a passive data capture followed by later processing).

A surprising outcome was that may firms decided to outsource certificates to companies like CloudFlare. This is great for the CEO who no longer gets woken up in the middle of the night with things like Heartbleed; now it's CloudFlare's problem. Unfortunately data may be less secure: encryption now runs from customer to CloudFlare, but not necessarily from CloudFlare to company actual servers unless a premium option is purchased. Of course companies don't pay the premium. So now data is backhauled across the Internet where it can be read with passive taps.



This is a logic bug in the implementation of the protocol. The failure here occurs because the client (not the server) gets to choose how many bytes (*x*) to return, so a malicious client can choose to return zero bytes. Further reading: https://en.wikipedia.org/wiki/Intel_Active_Management_Technology

Concurrency bug: time of check to time of use failure (TOCTOU)



In this example, a programmer is writing a program which has the setuid bit set (see Operating Systems course from last term). Therefore the programmer first checks whether the user has access to a particular file, then if true, uses the file by writing some data to it.

The bug occurs if the operating system can be coerced into performing a context switch at the red line, during which time a malicious user then (e.g. by updating symbolic links) swaps the file accessed. Then sometime later the program will write to a file which the user has specified which the user may not have write access to. This is called a *race condition*. We will see another race condition bug later in the course (Therac-25). Note that the concurrency here is not within the program itself, which has only a single thread of execution. Rather it occurs because the operating system supports multiple concurrent processes in execution.

Clallam Bay Jail inmates perform code injection on payphones

- 1. Inmate typed in the number they wished to call
- 2. Inmate selected whether the recipient spoke Spanish or English
- 3. Inmate was asked to say their name; "Eve", say
- 4. The phone then dialled the number and read out a recorded message in chosen language and appended inmate name to the end:

"An inmate from Clallam Jail wishes to speak with you. Press three to accept the collect call charges. The inmate's name is" ... "Eve"

The hack is to select the language that the callee *doesn't* speak (e.g. Spanish), and then state your name in Step 3 as "To hear this message in English press three".

Lesson: remember that you are protecting the whole system, including against potentially malicious users.