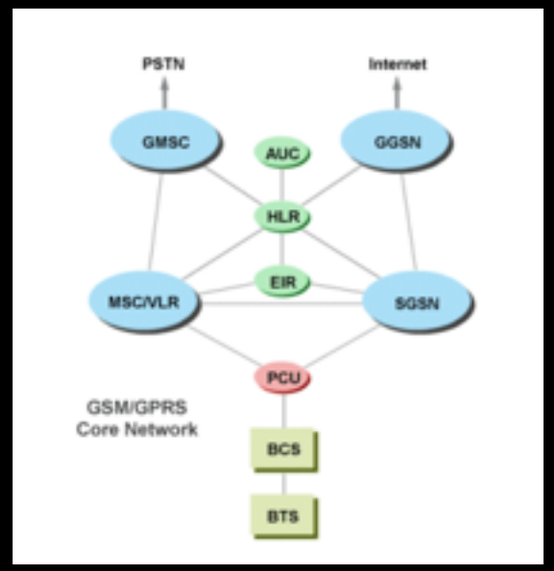
Long Term Exploitation

"Baseband security? 4Get about it."

Background: 2G

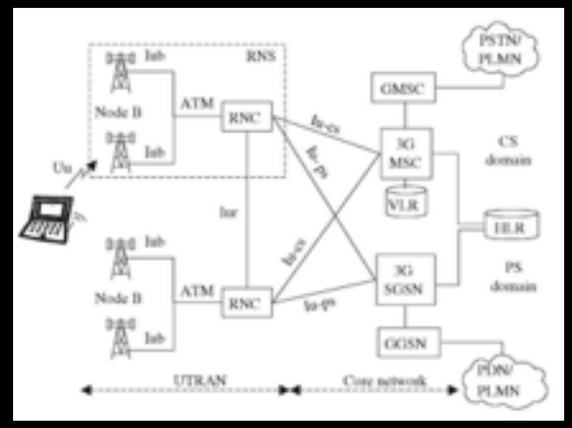
- GSM specification started in 1982
- Standardized by GSMA
- First commercial launch 1992
- TDMA based, circuit-switched
- 2.5G: GPRS (packet-switched) added in 2000



The graphics should say BSC not BCS

Background: 3G UMTS

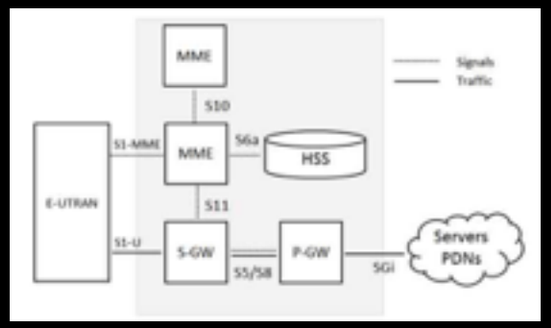
- 3GPP standard organization formed, UMTS/WCDMA started in 2000
- TDMA and CDMA variants, new Layer 1&2
- Same core network as 2G
- Still Circuit-switched & Packet-switched hybrid

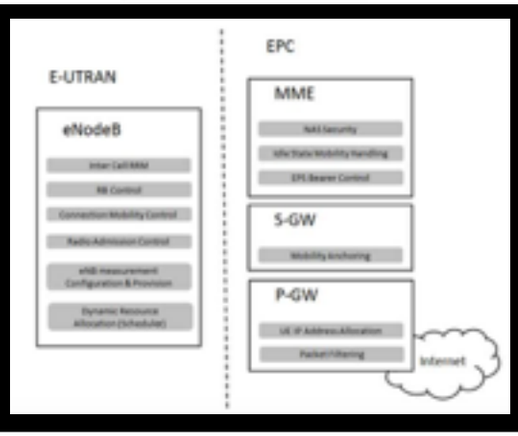


3GPP today standardizes 2/3/4G

Background: 4G LTE

- LTE specification started in 2004, Release 8 finalized in 2008
- First commercial launch 2010
- TDD and FDD
- “Simplified” network, all-IP architecture. Even calls are over IP (VoLTE)
- Higher bandwidth and lower latency, QoS support
- Fallback support for circuit-switched calls
- Note: LTE is in constant change, Rel13 is the currently ongoing release.





“E” actually stands for Evolved, not Exploitation. LTE is basically “slap the letter E in front of things”.

The “air interface” is the E-UTRAN (Evolved Universal Terrestrial Radio Access Network). The network elements that are the internals of the LTE network together form the EPC (Evolved Packet Core).

The cell tower is the eNodeB, it connects the User Equipment (UE, aka the phone) to the network.

There is control plane (CP; setup of things) and user plane (UP; actual traffic). Traffic goes over (logical) Radio Bearers (RB), Signaling and Data RBs for CP and UP, respectively.

In 2G/3G there was both Circuit-Switched (calls, SMS) and Packet-switched (data) traffic, but in LTE everything is Packet-switched.

Radio communication is the “Access Stratum (AS)”. Top-layer control protocol for it in LTE is Radio Resources Control (RRC), terminating in the eNB.

The functional layer between the phone and the core network is the “Non-Access Stratum (NAS)”. Top-level control protocol for it in LTE are EMM and ESM, terminating in the MME (Mobility Management Entity).

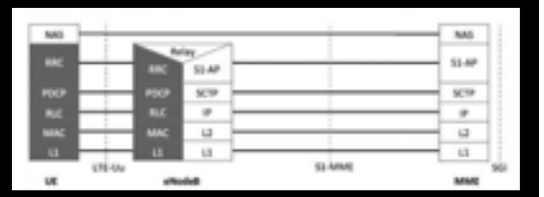
2G/3G NAS had Mobility Management (MM: movement, authentication, …) and Connection Management (CM: calls/SMS/SS) for Circuit-switched and GPRS Mobility Management (GMM) and Session Management (SM) for Packet-switched.

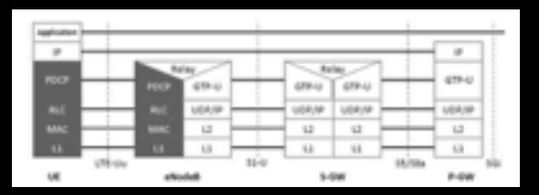
2G/3G to 4G Essentials

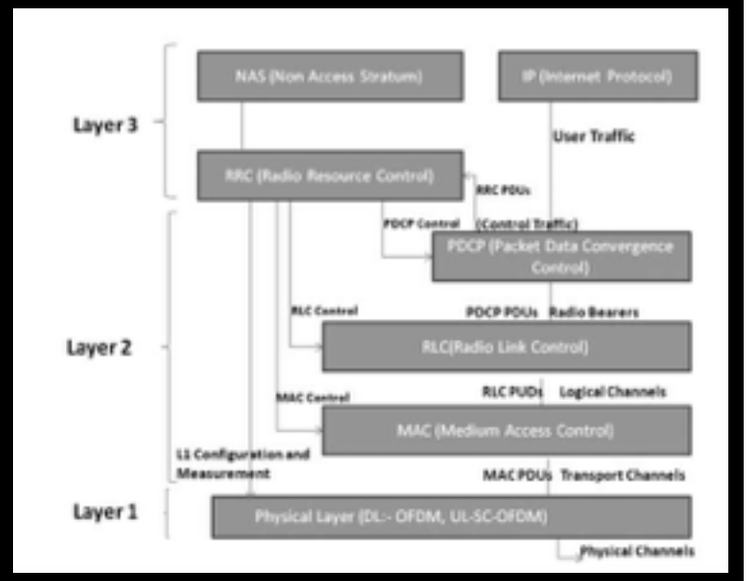
	2G/3G	LTE
Network Elements	GERAN and UTRAN	E-UTRAN
	BTS/BSC (GSM), NB/RNC (UMTS)	eNB
	SGSN/PDSN-FA	S-GW
	GGSN/PDSN-HA	PDN-GW
	HLR/AuC	HSS/AuC
	VLR	MME
Core Network Protocols	SS7-MAP/RADIUS	Diameter
	GTP v0/v1	GTP v2
	MIP	PMIP
AS Protocols	PHY/LAPDm/RR (GSM) PHY/MAC/RLC/RRC (UMTS CP) PHY/MAC/RLC/PDCP (UMTS PS UP)	PHY/MAC/RLC/PDCP/RRC
NAS Protocols	MM, CM (CS CP) GMM, SM (PS CP) IP (PS UP)	n/a (CS) EMM, ESM (PS) IP (PS UP)
Calls	Circuit-switched, controlled by Call Control in NAS CM	VoIP; CS Fallback*
SMS	Circuit-switched, controlled by SMS in NAS CM	SMS over IP; SG-SMS over NAS*
Suppl. Services (e.g. USSD)	Circuit-switched, conntrolled by Supplementary Services in NAS CM	Multimedia Telephony (IP); CS Fallback*

*transition solutions

LTE Protocol Stacks







What We Won't Talk About

- EPC internals
- VoLTE
- Handovers
- Circuit-Switched Fallback

The EPC is a beast of its own. Plenty of signaling protocols (S*, GTP, Diameter, etc.) for plenty of connections: eNB to MME, eNB relays, MME to MME, eNB to S-GW, S-GW to P-GW, etc. For the purposes of this talk, we "collapse" the EPC into the MME.

VoLTE works over IMS (IP Multimedia Subsystem). This is also a massively complicated architecture. For the purposes of this talk, we "collapse" the IMS into "IP traffic".

Handover is always network controlled in LTE, i.e. it requires an authenticated MME.

CS Fallback has many scenarios. Key derivation when transitioning between GERAN/UTRAN/E-UTRAN is non-trivial also, but not discussed here.

2G Security: Theory

- “Authenticity, Confidentiality, Privacy”
- User authentication based on per-subscriber secret key in SIM/AuC
- Stream ciphers to encrypt traffic on the air interface
  - A5/0 (null), A5/1, A5/2, A5/3 (KASUMI), A5/4 (KASUMI)
  - Frame number used as input against replays
- Temporary Identifier (TMSI) to protect subscriber privacy

Authentication and Key Derivation is based on a challenge-response algorithm and (originally) secret key derivation functions (COMP128-1/2/3/4).

3G Security: Theory

- Adds mutual authentication of the UE and NB
- Replaces the SIM with USIM (still compatible with SIM)
- Ciphering extended to NB-RNC link
- New ciphers, separate encryption and integrity
  - UEA0 (null), UEA1 (KASUMI), UEA2 (SNOW3G)
  - UIA0 (null), UIA1 (KASUMI), UIA2 (SNOW3G)
  - COUNTers used as input against replays

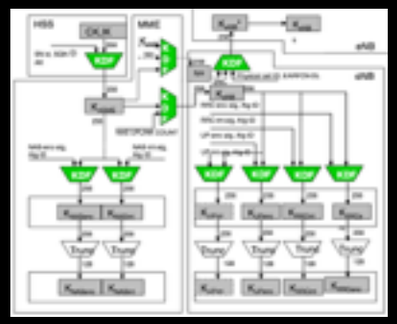
4G Security: Theory

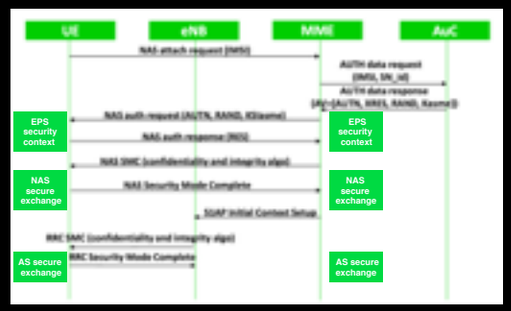
- Only USIM compatible
- New ciphers:
  - EEA0 (null), EEA1 (SNOW3G), EEA2 (AES), EEA3 (ZUK)
  - EIA0 (null), EIA1 (SNOW3G), EIA2 (AES), EIA3 (ZUK)
- Radio network (AS) and core network (NAS) security is separated
  - 2 layers of ciphering; AS terminates in eNB, NAS terminates in MME
- GUTI (~TMSI) to protect subscriber privacy
- IMEI ciphered to protect user equipment privacy

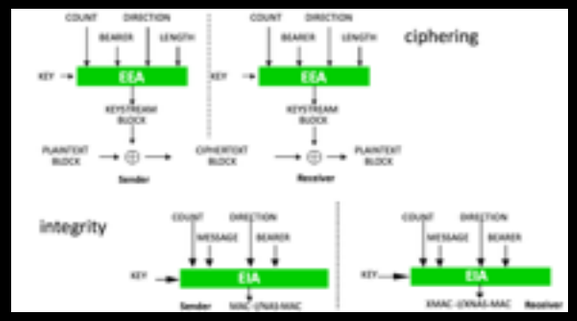
The AS ciphering implementation moved from MAC/RLC to PDCP layer. This was done as PDCP became a layer used both in CP and UP, a change from UMTS.

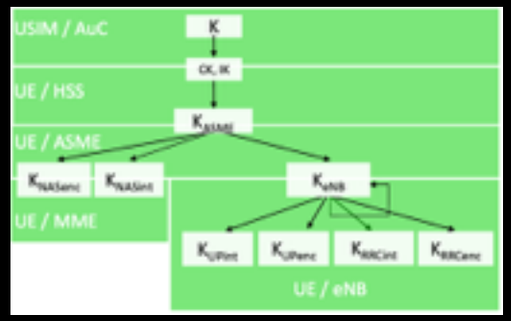
4G security also makes security changes inside the EPC (network domain security) and also adds complexity in the network with Relay Nodes. This is outside our scope.

4G Security: Theory









Key derivation and ciphering algorithms and the AKA procedure.

Sidebar: Lawful Intercept

- Lawful Intercept is supported in all of 2/3/4G
- Yes, network operators enable local authorities to silently locate, track, and intercept the communications of subscribers.
- A nice topic for debate, but entirely orthogonal to this presentation. We put this aside.

Attack Scenarios

Attack	Description
Impersonation	Stealing subscriber identities aka SIM cloning
Eavesdropping	Capturing & retrieving plaintext communication
Location Tracking	Tracking the movement of a subscriber through the network. Finding the precise location of a subscriber within a location/tracking area.
Identification	Finding out the identity of a UE (IMEI) or SIM (IMSI) connected to the network.
Man-in-the-Middle	Actively intercepting/modifying traffic.
Baseband Vulnerabilities	Exploiting implementation vulnerabilities in Layer2/3
Application Layer Exploitation	Exploiting vulnerabilities or insecure features in the application layer (e.g. Binary SMS).
Denial-of-Service	Attacks that cause permanent or temporary DoS to subscribers.
Core Network Attacks*	Targeting the core network directly.

*No research was done on core network attacks in LTE, this will not be discussed here.

Attacks on LTE

- With cipher and USIM improvements, there are no known ways to actually break the crypto, either to recover the K from the SIM, or to break the authentication, encryption or integrity protection.
- With two-way authentication, we can't impersonate eNBs either.
- So the common perception is that both passive and active attacks are thwarted in LTE.
- However, the reality is more complicated for 3 major reasons.

Attacks on LTE

- Not everything is encrypted
- The specifications allow for several messages without integrity protection
- Femto Cells: if one is compromised (by any physical or remote attack), AS security is compromised.

Some protocol layers and some messages simply aren’t encrypted. And in some network configurations nothing is encrypted. This can lead to Eavesdropping.

Since eNBs can send certain messages without integrity protection that UEs accept, malicious eNBs are partially possible. This can lead to Identification, Location Tracking, Network/Application Layer Exploitation, and Replay Attacks.

Compromising femtocells can enable interception/mitm as well. Note the 2 layer architecture of LTE security: this still does not compromise the NAS-layer communication with the MME (e.g. SG-SMS).

Attacks Enabled by Lack of Encryption

Eavesdropping

- Null encryption is supported for both AS (UP & CP) and NAS. IFF the network configures EA0, then the data is simply plaintext.
- How typical that is, hard to say. Maybe widespread, maybe extremely rare.
- On paper, Ciphering Indicators were mandated by the GSM 02.07. specification, but that specification also allows for the SIM to turn this off.
- In practice, mobile OSes do not provide this info.

Relevant specs: 33.401 5.1.3.2

Location Tracking #1: Presence Detection

- Scenario: verify whether a subscriber is in a tracking area or not.
- MAC provides different Logical Channels for different tasks: BCCH (broadcast), PCCH (paging), CCCH (common control), DCCH (dedicated control), DTCH (data traffic), etc.
- Broadcast* and Paging channels are never encrypted.
- If we trigger paging for a subscriber, we can observe and correlate pages to verify whether a subscriber is present in an area or not.
- This only works easily if the network pages by IMSI. If it pages by GUTI, an attack is still plausible, but a lot more difficult.

Paging can be triggered with a ring but also with e.g. Type0 SMSes. That is harder to do if the operator filters such SMSes in its network.

LTE Multicast (MBMS) has its own channels - MCCH, MTCH.

Relevant specs: 33.401. 7.4.1, 7.4.2 and 36.331. 4.2.2

Attacks Enabled by Lack of Integrity Protection

Null Integrity

- Both NAS and AS includes EIA0. If this is supported by the UE, all bets are off.
- Normally, EIA0 is only allowed for emergency calls.
- However, in early stages of LTE deployment, EIA0 creeped back in (again with the “transition”).
- Predictably.. baseband vendor code in 2014 still accepted EIA0. Found and disclosed by Benoit Michau (SSTIC 2014).

Integrity protection is also what protects against downgrade attacks: initial RRC connection request from the UE will encapsulate the UE’s ciphering capabilities, which is played back to the UE by the MME once security has been established. This is what tells the UE that the MME selected ciphers based on the UE’s true advertised capabilities.

Relevant specs: 33.401 7.2.4, 15.2.2

Access Stratum Integrity

- Nothing below PDCP SDUs are protected.
- Broadcast System Information (BCCH) and Paging (PCCH) is never protected.
- SRB0 (CCCH) is never protected.
  - RRC Connection Setup, Reject, Re-establishment Reject
- SRB1 (DCCH) is only protected after “AS security has been activated”.
- SRB2 (DCCH) is always protected.
  - Downlink Information Transfer (NAS messages)
- DRBs (DTCH) are never protected: there is only encryption in User Plane, no integrity protection.

“Below PDCP SDU” meaning PDCP PDU headers and the layers below: MAC, RLC.

Access Stratum Integrity

- The SRB1 case is more complicated.
- Messages allowed “after AS security has been activated”:
  - Handover, Connection Re-configuration for handover or security, Relay Node Configuration, SMC
- Other messages:
  - UE Capability Inquiry, Connection Reconfiguration for Measurements, DL Information Transfer, Counter Check, Connection Release

Establishing what messages on SRB1 (CCCH) are allowed without integrity protection is quite difficult based on the specs.

33.401 says that 36.331 is supposed to list exceptions explicitly, but it does not. An explicit exception is not even listed where it is obviously imperative (Information Transfer that encapsulates NAS messages).

Only allowed after “AS security has been activated” or must “pass the integrity protection check” is listed for Handover, Connection Re-configuration (for handover and security), Relay Node Configuration, and Security Mode commands and that is it, no other mentions.

It is derived from this that the other possible CCCH messages would be accepted without integrity protection.

Relevant specs: 33.401 5.1.4.1, 36.331. 4.2.2, 5.2.1, 5.3.1.1, 5.3.1.2, 5.3.4.3, 5.3.5, 5.4.3, 5.4.4

Identification

- Scenario: fingerprinting for exploitation. Identify the user equipment / baseband version of a subscriber.
- Run UE Capability Inquiry.
- In total, there are more than 120 capability fields.
- If sufficiently unique, capabilities may be usable to identify the type of equipment that a subscriber has.

Relevant specs: 36.331. 6.3.6

Location Tracking #2: Precise Location

- Scenario: identify precise location of a subscriber.
- Configure the UE to perform measurements.
- Measurement reports may be usable to identify a more precise location of the UE.

Relevant specs: 36.331. 5.3.5, 5.5

User Plane Replay Protection

- User plane encryption uses a COUNT for replay protection.
- Unless EEA0 is used, any modification/injection/ replay of user plane data results in garbage.
- So normally, we could only alter LTE user plane traffic with a compromised femtocell.
- However, there is a loophole in the specification that enables user plane message replays.

Note for cryptographers: whether packet modification is feasible or not in a meaningful way actually depends on the “malleability” property of the encryption cipher in use, such as SNOW3G. This should be studied carefully. Out of scope here.

User Plane Replay Protection

- COUNT is made up by concatenating the SN (sequence number) and the HFN (hyperframe number).
- UE keeps track of the next expected SN for both RX and TX.
- Only the SN is sent in a PDCP PDU. The HFN is maintained locally by both the UE and the eNB.

User Plane Replay Protection

1. If SN < Next_SN: HFN += 1
2. Decipher message using COUNT := HFN|SN
3. NEXT_SN := SN + 1
4. If NEXT_SN > MAX_SN: NEXT_SN := 0; HFN += 1
5. Decompress message
6. If message is erroneous, discard
7. Deliver to upper layer

The point is that if we get a message we already saw, it will either have a smaller SN, in which case HFN is incremented before deciphering, so the result will be garbage, or it will have the next or larger SN, but in that case that message is from the future (i.e. has not been sent yet actually so can't be replayed) or it was sent way before with a previous HFN, which means that with the current HFN the deciphering will also result in garbage.

This is what the attack circumvents by wrapping the HFN around to current HFN-1.

Note that this is the algorithm for PDCP Unacknowledged Mode (UM). Acknowledged Mode (AM) is more complicated, it uses TCP-style sliding windows. Nonetheless, the same attack described for UM applies.

User Plane Replay Attack

- The attack is based on overflowing the HFN of the UE.
- The specification does not mandate any action by the UE for HFN overflows.
- It only says that the eNB must prevent this from happening, but that assumes a benign use-case.

Regarding the practicality of this attack.

On paper, RRC Counter Check procedure detects out-of-sync COUNTs, however the spec does not mandate that the UE trigger that if the HFN overflows, or at any other time. Regardless, the specification also does not ban the UE from doing that, so this will be heavily implementation dependent.

Also note that practicality depends on the type of UP traffic and whether there's replay protection applied in higher layers.

Relevant spec: 36.323 5.1.2.1.3

User Plane Replay Attack

1. Send message with MAX_SN
    1. SN < Next_PDCP_SN is False
    2. Message is deciphered
    3. NEXT_SN = SN + 1 = MAX_SN + 1 = 0; HFN += 1
    4. Message is discarded
2. Repeat 1. until HFN wraps around and becomes HFN_old - 1
3. Send message with MAX_SN - 1
    1. SN < Next_SN is False
    2. Message is deciphered
    3. NEXT_SN = SN + 1 = MAX_SN
    4. Message is discarded
4. Now replay the message with SN_old
    1. SN_old < NEXT_SN is True, HFN = HFN + 1 = HFN_old
    2. Message is deciphered correctly with HFN_old and SN_old

Application Layer Exploitation

- Broadcast SMS services (CMAS, ETWS) are delivered via System Information messages in the BCCH
- BCCH has no encryption, no integrity protection
- Not very assuring that these can be injected or sent by a malicious eNB

“Broadcast SMS”’ proper name is Cell Broadcast Service.

CMAS is Commercial Mobile Alert Service, it is delivered in SystemInformationBlock Type 12

ETWS is Earthquake and Tsunami Warning Service, it is delivered in SystemInformationBlock Type 10 & 11

Relevant specs: 36.331 5.2.2

Baseband Vulnerabilities: AS Attack Surface

- All the PDU parsing in lower layers (MAC, RLC, PDCP)
- ROHC decompression in PDCP if EA0 is used
- System Information parsing (e.g. ETWS, CMAS)
- ASN1 decoding in RRC
- Processing of all the RRC messages accepted without integrity protection
- Implementation of the logic for when to discard messages without valid integrity protection

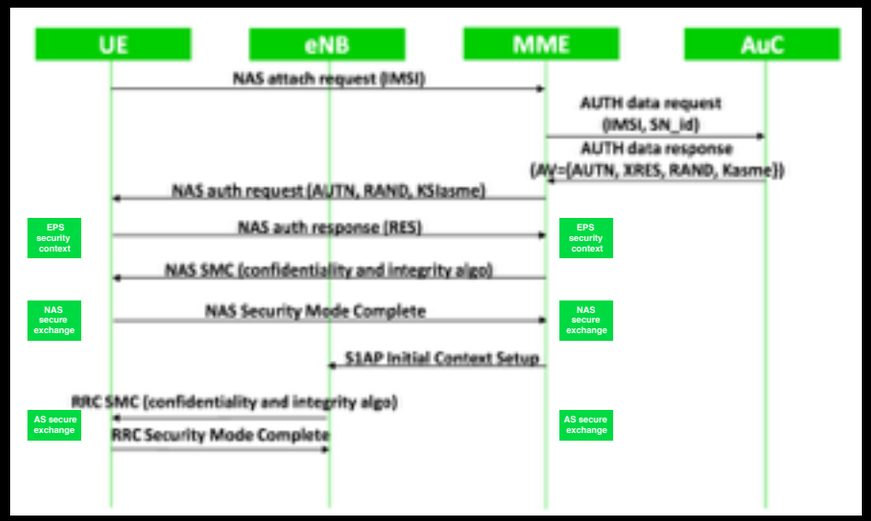
Every message in RRC is ASN1 encoded. Since many RRC messages are accepted without integrity protection, this effectively exposes the ASN1 decoder library of implementations to malicious input.

PDCP's two main functions are AS-level ciphering and compression. RObust Header Compression is used to save bandwith by compressing IP packets. An incoming message in the UE is first deciphered and then decompressed; that is why the attack surface is exposed mostly when EA0 is used. It is always exposed to a compromised Femtocell.

NAS Integrity

- For initial attachment, UE always has to establish an EPS security context. This is done via the AKA (Authentication and Agreement).
- The UE's initial NAS message triggers the AKA.
- The secure exchange of NAS and AS messages is established after the AKA via the NAS and AS SMC procedures, respectively.

NAS Integrity



AKA, NAS SMC, and AS SMC in one picture.

First, the UE establishes an RRC connection, to go from RRC-IDLE to RRC-CONNECTED. When it enters RRC-CONNECTED, it also automatically goes from EMM-IDLE to EMM-CONNECTED. (This is not shown on the diagram.)

NAS Integrity

- The specification lists the messages that are accepted before the secure exchange of NAS messages is established.
- “These messages are accepted by the UE without integrity protection, as in certain situations they are sent by the network before security can be activated.” - 3GPP 24.301
- By itself, the window from power-on to establishing the EPS security context initially is not practical to exploit in most scenarios.

Clearly, these can be sent when a UE has not yet established an EPS security context. In fact, the EPS security context has to survive even switch-offs of the device. (The specification describes how it must be saved in non-volatile storage)

List of messages:

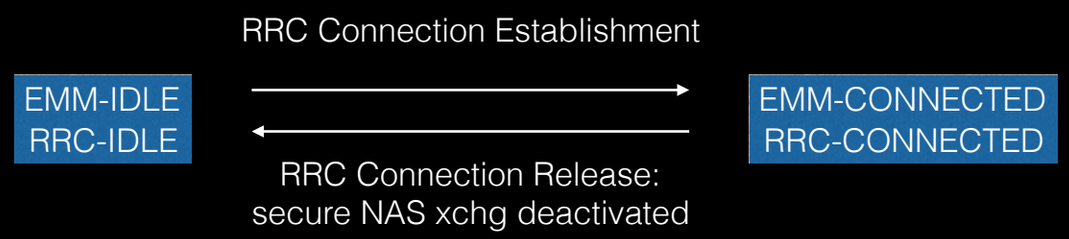
Identity Request for IMSI
Authentication Request
Authentication Reject
Attach Reject (not cause #25)
Detach Accept (non switch-off)
Tracking Area Update Reject (not cause #25)
Service Reject (not cause #25)

Relevant spec: 24.301 4.4.4.2

NAS Integrity

- What happens after the initial attachment?
- The UE moves around, becomes idle, uses services again, changes eNBs, etc.
- How do transitions between different states occur?
- How do these transitions effect the security state?

NAS Integrity



- RRC connections are released by the network due to user inactivity.
- The NAS connection is automatically released in this case, terminating the “secure exchange of NAS messages”.
- User inactivity is determined by implementation-specific IDLE timers. In practice, this is on the order of seconds.

The phone starts are RRC-IDLE/EMM-IDLE. When initial RRC connection is established, it goes to RRC-CONNECTED. By the specification, it automatically enters EMM-CONNECTED if it enters RRC-CONNECTED.

Once in EMM-CONNECTED during the initial attachment, the secure exchange of NAS messages is established with the SMC following the AKA, as described before.

NAS connection is released if the RRC connection is released (RRC-CONNECTED -> RRC-IDLE).

RRC connection may be released by the eNB or MME due to misc. reasons, including user inactivity, RRC signaling integrity check failures, other network failures, etc.

Most important is the fact that user inactivity triggers this as well. Note that an active attacker can likely inject traffic that will result in errors that will trigger the network to release the connection as well.

Note: 23.401 uses the terms ECM-IDLE, ECM-CONNECTED; 24.301 uses EMM-IDLE, EMM-CONNECTED and its abbreviations section explains that the two (ECM/EMM) refer to the same.

Relevant specs: 24.301 5.3.1.2, 5.3.1.1; 36.331. 5.3.

NAS Integrity

- In LTE, UE mobility (handover) is controlled by the network when in CONNECTED states.
- In IDLE state, the UE will perform cell re-selection at its own discretion, similar to 2G/3G.

NAS Integrity

- At some point, the UE can decide to use services, maybe because it receives a paging event.
- Now, the “secure exchange of NAS messages” has to be re-established.
- The UE sends its initial NAS message integrity protected but not ciphered and includes a key identifier (eKSI) with the message.
- However, after sending this message the UE still does not consider the security re-established.

NAS Integrity

- A UE with an active EPS security context considers the secure exchange of NAS messages re-established only if:
  - eNB responds to the initial NAS message with a NAS SMC,
  - eNB responds to the initial NAS message with an integrity protected NAS response,
  - eNB responds with an integrity protected RRC Connection Reconfiguration, which re-activates AS security and then the user plane data bearers are re-established.
- Until then, the UE will process (other) messages in the state of “secure exchange of NAS messages not established”!

Relevant specs:
24.301 4.4.2.3

LTE IMSI Catcher

- Scenario: send an Identity Request for IMSI to collect subscriber identity.
- Important to note that this truly is an “IMSI Catcher” and not more.
- The attack does not enable Location Services tricks, because LTE LPP goes over user plane (SUPL) or NAS.
- It also does not enable interception as user plane data traffic does not proceed until AS security has been re-activated.

LTE DoS

- Scenario: send an Attach/Tracking Area Update (TAU)/Service Reject to cause Denial-of-Service.
- Reject messages can have various causes.
- Some reject causes (#3, #6, #8) disable the USIM entirely until power-off or USIM removal.
- Many causes trigger the UE deleting its GUTI and eKSI:
  - without a GUTI, the UE will use IMSI in follow-up requests
  - deleting eKSI is analogous to removing the EPS security context

Tracking Area is analogous to Location Area in 2G.

Normally, the EPS security context is not expected to be flushed. However, the UE uses the eKSI to tell the network in a NAS request with Kasme to use. If the UE deletes all its eKSIs, it can’t include it in an initial NAS request, which practically means the associated EPS security context won’t be taken into use anymore.

Relevant specs: 24.301 5.5.1.2.5, 5.5.1.3.5, 5.5.3.2.5, 5.6.1.5

LTE DoS via MME

- On the MME side, the specification similarly lists the messages that will be accepted without integrity protection.
- “The (…) messages are accepted by the MME without integrity protection, as in certain situations they are sent by the UE before security can be activated. “
- This includes Attach and Detach Requests.
- Consequently, the classic DoS attacks (“IMSI Detach” and “IMSI Flood”) can be plausible in LTE as well.

Relevant spec: 24.301 4.4.4.3

LTE DoS to Downgrade

- Scenario: send an Attach/TAU/Service Reject to downgrade to 2G/3G.
- Reject cause #7 for Attach Requests disables the USIM for EPS until power-off or USIM removal.
- Reject causes #7 and #14 for periodic TAU Requests will trigger the UE to attempt GERAN/UTRAN instead.
- The same is true for Service Requests.
- Note: jamming is a viable alternative anyway.

This behavior for Reject causes #7 and #14 for TAU requests only applies if the UE is operating in CS/PS mode.

Service Reject cause #18 is exactly “CS Fallback not available”. This is specifically meant for telling the UE responding to a page on a non-VoLTE network that CS Fallback is not supported.

Note: in practice one would have to consider what mode targets operate in (CS/PS 1, CS/PS 2, PS only).

Also note that current commodity mobile OSes do not allow users to turn off RAT support (e.g. GERAN/UTRAN) entirely. Preference selection options exist, but these do not prevent fallbacks entirely.

Relevant specs: 24.301 5.5.1.2.5, 5.5.1.3.5, 5.5.3.2.5, 5.6.1.5

Baseband Vulnerabilities: NAS Attack Surface

- Parsing of NAS messages accepted without integrity protection.
- Implementation of the logic for when to discard messages without valid integrity protection.
- Example, found and disclosed by Benoit Michau (Qualcomm Security Summit 2015):
  - Due to a state-machine bug, the baseband vendor's implementation always accepted detach messages without integrity protection, regardless of NAS security state.

NAS (EMM) message encoding is all TLV encoded Information Elements, just like MM in 2G.

Relevant specs: 24.301

Thank you!

questions: @kutyacica

Appendix: History of 2G/3G attacks

Impersonation

- In GSM, authentication responses (SRES) and session keys (Kc) are generated by the SIM from the key K with an algorithm called COMP128.
- COMP128-1/2/3 were secret algorithms.
- COMP128-1 turned out to be completely broken. Reverse engineered in 1997, broken in 1998. SIM cards based on it could easily be cloned.
- COMP128-2 is more secure, but only provides 54 bits of entropy for the session key. COMP128-3 provides the full 64 bits. Both were secret and reverse-engineered.
- COMP128-4 uses Milenage. This is the 3GPP standard algorithm based on AES used by USIMs in UMTS/LTE.
- Since the days of COMP128-1, no direct SIM card cloning attacks are known.

Impersonation

- Fast-forward to 2013.
- SIM cards run Java applets. (What? Yes.)
- These can be provisioned Over-The-Air (OTA), via SMS, directly delivered to the SIM.
- AES & 3DES can be used, but many SIM cards still used DES to protect these OTA updates.
- SRLabs researchers showed a way to break this application of DES in order to install malicious Java applets on SIM cards.
- Even worse, the Java sandboxes of some SIM vendors were weak. Breaking out of the sandbox gets the “keys to the kingdom”, i.e. remote SIM cloning.
- Today, absent such “old” SIM cards using DES, no attacks are known.

Impersonation

- Even if SIM cards are not cloned, session keys can be re-used as long as they are valid.
- Depending on network operator practices for longevity of session keys, this attack may be practical or impractical.
- Above all, it requires the ability to crack session keys, which takes us to …

Eavesdropping

- A5/2 was broken by Goldberg and Wagner the month it was published. It can be trivially cracked in real-time. Today it is banned.
- A5/1 died a slow death.
  - Several known-plaintext based attacks using pre-computation were published in 2000 (Biryukov et al, Birham et al).
  - Finally, in 2009, the A5/1 Cracking Project demonstrated the attack in practice and released pre-computed rainbow tables that enable breaking a session key in seconds.
  - Note: because it is a known-plaintext attack, a padding randomization “patch” for A5/1 exists

Impersonation +Eavesdropping

- Paging in GSM is essentially a race.
- The network pages in plaintext and whichever UE responds first is “believed” to be the correct UE.
- This makes it possible to hijack calls/SMS by winning the race.
- If the network authenticates the UE, then this is a DoS only. (It can apply to entire location areas as well!)
- However, GSM does not mandate the network to authenticate every single time. IF the network does not authenticate, then calls/SMS can actually be hijacked.
- What’s worse, if the session key is breakable, then even with ciphering, the plaintext can be acquired.
- Nico Golde demonstrated these attacks against GSM in practice (29C3, 2012).

Identification

- Lazy use of TMSIs is one of the constant privacy issues with operator networks.
- One recorded example: E-Plus only using IMSI in 2012 (reported by Nico Golde).
- The GSMMAP project has a lot of data covering network operator practices of all three mentioned aspects: impersonation, eavesdropping, identification.

IMSI Catchers

- Lack of mutual authentication in GSM allows for rogue base stations.
- Often labeled “IMSI Catchers”, which is something of a misnomer when capabilities include full MitM.
- May also use Control Plane Location Services (RRLP in GSM/UMTS) to precisely locate a subscriber.
- Harald Welte in 2009 showed that RRLP can be leveraged this way.
- First public demonstration of interception by Chris Paget at DefCon18 (2010).
- Many (secretive) commercial solutions exist.

Baseband Vulnerabilities

- SMS User Data Header parsing vulnerabilities
  - Traditionally (feature phones) SMS parsing was implemented in the baseband, in modern phones the mobile OS parses the SMS PDU
  - SMS Fuzzing Android&iPhone: Miller and Mulliner (Black Hat 2009)
  - SMS Fuzzing Feature Phones: Golde and Mulliner (27C3 2010)
- NAS TLV parsing vulnerabilities
  - Ralf-Philipp Weinmann found RCE vulnerabilities in Infineon and Qualcomm basebands. RCE via BoF in the authentication challenge was publicly demonstrated (DeepSec 2010).
  - In 2013, GSMK announced that they have found RCE vulnerabilities in Infineon and Qualcomm basebands, but the details were never disclosed.
  - Benoit Michaut (SSTIC 2014, vendor/details not disclosed).

Baseband Vulnerabilities

- Sidebar: local attacks against basebands
- Not a remote threat. Technically, chained with an initial RCE, compromising the baseband from the mobile OS enables “temporary impersonation” by stealing session keys
- Geohot et al. iPhone Infineon baseband AT command parsing vulnerabilities (2009)
- Guillame Dellugre Qualcomm baseband DIAG commands to peak/poke memory (28C3 2011)
- Marcus Vervier MediaTek baseband firmware modification from Android (HITB 2015)

Application Layer Exploitation: SMS

- Not all SMSes are meant for and seen by the user.
- Type0 SMS (“silent SMS”) is silently acknowledged. This can be leveraged for tracking, since it’s a paging event that the UE silently and automatically responds to.
- Binary SMS is used to deliver “SIM Toolkit Application” (STK) messages. Not seen by the subscriber.
- WAP Push SMS is used to initiate an “OMA-DM” (Open Mobile Alliance Device Management) connection. Also not seen by the subscriber.
- If a network operator does not filter these types of SMSes in their network, then attacks are feasible without base stations too.
- In past years (~<2010), open Internet services for sending arbitrary SMSes (including source spoofing) were wildly available. Operators typically filter these today.

Application Layer Exploitation: STK

- STK messages are passed directly to the SIM card by the baseband.
- SRLabs SIM cracking attack described previously (Black Hat 2013)
- Gordeychik et al. (PacSec 2014) also showed that some STK applications (“TAMs”) on some SIMs respond to commands without authentication.
  - In some cases, a file system TAM on a SIM could simply be queried for the Kc.
  - Such insecure configuration is “rare” (not well quantified). But with billions of active SIM cards, even a small percentage is a significant number.

Application Layer Exploitation: OMA-DM

- WAP Push SMS for OMA-DM is passed from the baseband to the mobile OS, typically silently consumed by a dedicated process.
- OMA-DM is used for OTA provisioning and updates. The client responds to the initial WAP Push message by connecting to the DM server requested.
- In 2009, MSEC researchers demonstrated that by spoofing OMA provisioning SMSes, mobile data connections could be hijacked (reconfigure DNS etc). This was the old days of no authentication, no operator filtering, and the ability to simply send arbitrary SMSes using an Internet service (Black Hat 2009).
- Solnik and Blanchou demonstrated that the weak OMA-DM authentication in Redbend's implementation can be exploited (Black Hat 2014).
    - TLS is optional in OMA-DM. Instead, there is password based authentication and HMAC-MD5. The username is the IMEI/MEID. Solnik and Blanchou found that passwords were static across all devices per Carrier.
    - Worse, the researchers discovered multiple RCE vulnerabilities in the Redbend stack.
    - Redbend has a 70-90% market share (Android, iPhone, Blackberry, WP, 3G dongles, etc).

Application Layer Exploitation: SUPL

- SUPL (Secure User Plane Location Protocol) is a user plane implementation of Assisted GPS. Basically, the phone reveals its location to the SUPL server when requested.
- SUPL in some phones is implemented in the baseband, in other phones it is implemented in the mobile OS itself and works over Wifi too.
- Normally, SUPL is over trusted connections (HTTPS).
- Ralf-Philipp Weinmann discovered that several SUPL implementations (Android, Blackberry) configured insecure (HTTP) SUPL servers. He also found vulnerabilities in the SUPL stack of Qualcomm's baseband. (Black Hat 2012)
- Martin Sauter in 2014 disclosed that the SUPL implementation on Android sent the IMSI to supl.google.com and it did not validate certificates.

Application Layer Exploitation: USSD

- Borgaonkar showed that USSD codes can be automatically triggered via URI handlers on Samsung devices in order to wipe phones or “kill” SIM cards. (Ekoparty 2012)
- Could be triggered via a malicious website, but also via WAP Push SMS.

Application Layer Exploitation: MMS

- Something something something stagefright
- Everyone knows this now :)
- Joshua Drake, Black Hat 2015
- Brian Gorenc et al. also talked about MMS fuzzing at DefCon 22 (2014)

Femtocell Exploitation

- The THC published research on taking over Vodafone 3G Femtocells in 2011. Compromise via serial, mods to enable IMSI catching, interception.
- Nico Golde & Kevin Redon published attacks on SFR 3G femtocells at Black Hat 2011.
  - Exploit insecure firmware update protocol to takeover an SFR femtocell locally
  - From there take over any SFR femtocell remotely
- Tom Ritter et al. presented a hack of Verizon 3G Femtocells at Black Hat 2013, once again enabling interception and also cloning.
- Osipov and Zaitsev: compromise Huawei 3G femtocell with physical access (Black Hat 2015)

Core Network Attacks

- SS7 connects operators' networks. No authentication built-in.
- "Weakest link": any operator enabling outside access brings down the entire walled garden. Getting access without being an operator is not trivial, but feasible.
- In 2014, Tobias Engel and SRLabs both presented research at 31C3, showing that SS7 commands can be abused to:
  - get subscriber's location
  - remotely configure call-forwarding
  - query the session keys of subscribers (2G/3G)

Denial-of-Service

- With Radio Access Technologies, physical jamming is always a threat.
- Back in the day of open SMS relays, Traynor et al. discovered that SMSes can be easily used to consume random access channels and effectively DoS entire location areas remotely (CCS 2005).
- In 2009, Dieter Spaar showed that a local attacker can also starve a serving cell of RACHs with a flood of requests.
- Similarly, the HLR/VLR could be flooded with attach requests (“IMSI Flood”) to prevent network access to others (Grugq, Black Hat 2010).
- In 2010, Sylvain Munaut published that unauthenticated IMSI Detach messages sent to the network in the name of another subscriber cause a temporary DoS to a targeted IMSI.
- IMSI Catchers can also cause temporary (e.g. until phone reboot) DoS by sending a Location Update Reject message. (Domonkos Tomcsányi, Hacktivity 2014)