Khan, Mohsin; Ginzboorg, Philip; Järvinen, Kimmo; Niemi, Valtteri

Defeating the downgrade attack on identity privacy in 5G

Published in:
Security Standardisation Research - 4th International Conference, SSR 2018, Proceedings

DOI:
10.1007/978-3-030-04762-7_6

Published: 01/01/2018

Document Version
Peer reviewed version

Please cite the original version:
Khan, M., Ginzboorg, P., Järvinen, K., & Niemi, V. (2018). Defeating the downgrade attack on identity privacy in 5G. In C. Cremers, & A. Lehmann (Eds.), *Security Standardisation Research - 4th International Conference, SSR 2018, Proceedings* (pp. 95-119). (Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics); Vol. 11322 LNCS). https://doi.org/10.1007/978-3-030-04762-7_6

Defeating the Downgrade Attack on Identity Privacy in 5G

Mohsin Khan[1,2](✉), Philip Ginzboorg[3,4], Kimmo Järvinen[1,2], and Valtteri Niemi[1,2]

[1]University of Helsinki, Helsinki, Finland
[2]Helsinki Institute for Information Technology, Helsinki, Finland
{mohsin.khan, kimmo.u.jarvinen, valtteri.niemi}@helsinki.fi
[3] Huawei Technologies, Helsinki, Finland
[4] Aalto University, Espoo, Finland
philip.ginzboorg@huawei.com

Abstract. 3GPP Release 15, the first 5G standard, includes protection of user identity privacy against IMSI catchers. These protection mechanisms are based on public key encryption. Despite this protection, IMSI catching is still possible in LTE networks which opens the possibility of a downgrade attack on user identity privacy, where a fake LTE base station obtains the identity of a 5G user equipment. We propose (i) to use an existing pseudonym-based solution to protect user identity privacy of 5G user equipment against IMSI catchers in LTE and (ii) to include a mechanism for updating LTE pseudonyms in the public key encryption based 5G identity privacy procedure. The latter helps to recover from a loss of synchronization of LTE pseudonyms. Using this mechanism, pseudonyms in the user equipment and home network are automatically synchronized when the user equipment connects to 5G. Our mechanisms utilize existing LTE and 3GPP Release 15 messages and require modifications only in the user equipment and home network in order to provide identity privacy. Additionally, lawful interception requires minor patching in the serving network.

Keywords: 3GPP · IMSI catchers · Pseudonym · Identity privacy · 5G

1 Introduction

A generic mobile network has three main parts: (i) user equipment (UE); (ii) home network (HN), i.e. the mobile network where the user has a subscription; and (iii) serving network (SN), i.e. a mobile network that the UE connects to in order to avail services.[1] The UE includes mobile equipment (ME) and a universal integrated circuit card (UICC). The UE is said to be roaming when the SN and HN are different. When both the UE and HN are made to 5G specifications, it is possible that the 5G UE connects to a legacy serving network, e.g., the LTE

[1] The notation used in this paper is summarized in Appendix A.

(long term evolution, or 4G) serving network. It is important to allow such a "downgraded" connection, because – especially in the early phases of 5G adoption – 5G coverage will be spotty compared to LTE.

International mobile subscriber identity (IMSI) is a globally unique identity of a mobile user. Identity privacy means that long-term user identities remain unknown to everyone else besides UE, SN, and HN. IMSI catchers, i.e. malicious devices that steal IMSIs in order to track and monitor the mobile users, are a threat to users' identity privacy [1–4]. Passive IMSI catchers attack by eavesdropping; active IMSI catchers attack by impersonating an SN. Mobile networks had protection against passive IMSI catchers since GSM, but until 5G they did not protect users against active IMSI catchers.

Third generation partnership project (3GPP) Release 15, the first 5G standard, includes protection against active IMSI catchers [5]. This protection is implemented so that the UE encrypts its identity using public key encryption with the public key of the HN [6]. The concealed identity is called subscription concealed identifier (SUCI). This protection works only when the SN is also a 5G entity because an SN from LTE, 3G or GSM networks would not know how to process the SUCI. This implies that an active IMSI catcher can mount a downgrade attack against a 5G UE so that it impersonates an LTE SN and exploits LTE's weakness in order to steal the IMSI of the 5G UE. In this paper, we propose mechanisms to prevent this downgrade attack.

Our solution uses pseudonyms that have the same format as IMSI for LTE communication to defeat the downgrade attack. The idea of using pseudonyms in IMSI format to confound IMSI catchers in mobile networks has been studied in several works during the recent years [7–12]. We take a similar approach as proposed in these papers. A pseudonym looks like a normal IMSI, but its last nine to ten digits are randomized and frequently changing.[2] The UE is provisioned with two pseudonyms in the beginning and it gets fresh pseudonyms during AKA runs. It uses these pseudonyms instead of IMSI to identify itself when connecting to an LTE SN. The UE uses the SUCI instead of IMSI to connect to a 5G SN. Our solution piggybacks on existing messages involved in the LTE and 5G authentication and key agreement (AKA) protocols to deliver new pseudonyms to the UE and does not require additional messages.

Since the space of pseudonym having IMSI format is limited, pseudonyms need to be reused, i.e., disassociated from one user and reallocated to another. On the other hand, a 5G UE may connect to multiple SNs simultaneously [6]; causing the UE to use different pseudonyms to connect to different SNs. Hence, it is a challenge for the HN to know when it can disassociate a pseudonym from a current user and reallocate it to a new user. In our solution the UE embeds information about its LTE pseudonyms into the SUCI. This enables efficient

[2] The first five to six digits of the IMSI identify the country and the home network of the mobile user. Even though these digits allow linkability in certain cases, (e.g., if in a visited network there is only one roaming UE from a specific country), these digits are not randomized, because they are needed to route initial requests for authentication data for roaming UE to the correct home network.

reuse of pseudonyms in the HN and accurate billing. The UE does not need a pseudonym to connect to a 5G SN. Hence, even in the unlikely event, where the UE and HN lose synchronization of pseudonyms, the synchronization can be restored because the UE obtains a new LTE pseudonym from the HN simply by connecting to a 5G SN.

This paper advances protocol design, beyond what is described in papers [7–12], as follows:

(i) The HN can reuse pseudonyms even when a UE has simultaneous connections to multiple SNs.
(ii) The states of pseudonyms in UE and HN will remain synchronized, given that the HN and the UE function correctly.
(iii) For the case where UE and HN get desynchronized due to some unlikely errors, we have detailed a re-synchronization mechanism for pseudonyms' state in UE and HN. The mechanism is based on running a 5G AKA with a 5G SN.

Any enhancement of user identity privacy must support Lawful Interception (LI), i.e. selective interception of individual subscribers by legally authorized agencies. In Section 5.2 we describe how LI is supported in the Release 15 enhancement to user identity privacy (where UE encrypts its identity) and propose to supplement pseudonym-based solution with similar features. Adding support for LI into our pseudonym-based solution requires software update (patching) in the core network elements of the LTE SN; it does not require patching in the radio access network elements (base stations).

In this paper we consider the LTE SN in detail. Same solution can be adapted for 3G SNs in a straightforward manner. Its adaptation to GSM would require additional measures because GSM lacks means to authenticate any message received by a mobile device. An example of such a measure has been described in [7].

2 Preliminaries

An IMSI is a globally unique number, usually of 15 decimal digits, identifying a subscriber in GSM, UMTS and LTE systems. The first 3 digits of an IMSI represent the mobile country code (MCC); the second 2 or 3 digits represent the mobile network code (MNC); and the last 10 or 9 digits represent the mobile subscription identification number (MSIN) [13]. The long-term subscriber identifier in 5G system is called Subscription Permanent Identifier (SUPI) [14]. It may contain an IMSI, or a network-specific identifier for private networks.

Figure 1 illustrates a high-level architecture of mobile networks. The UE of a user, which has a subscription with the HN (home network), connects with the SN (serving network) to get services. If the SN and the HN are different networks, we say that the UE is roaming. In that case, the SN and the HN connect with each other over the IP Exchange (IPX) network. The link between the UE and the SN is initially unprotected in LTE, 3G and GSM networks. In

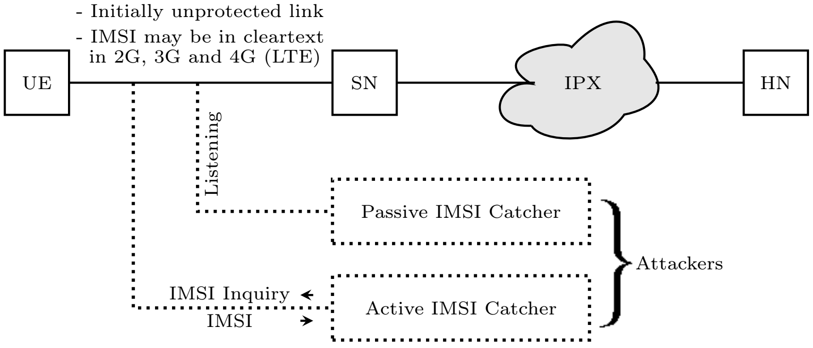

Fig. 1: Schematic illustration of mobile network.

5G, some information sent over this initial link may be confidentiality protected using the public key *pk* of the HN. Information needed for routing in IPX cannot be confidentiality protected (otherwise, routing would not work) [15].

The HN stores the (IMSI, K) pairs of all its subscribers in the subscription database. The 5G HN also has a public/private key pair *pk*, *sk*. The UE includes an ME and a UICC. The UICC is tamper resistant: a malicious entity can not read its content without sophisticated instruments. The universal subscriber identity module (USIM) is an application that runs in the UICC.

In LTE, the USIM stores the IMSI and the subscriber-specific master key K of the user. Please note that if a UE conceals the IMSI using K, the HN would not know which K to use to decrypt the message. Moreover, the UE is not provisioned with any SN-specific keys. Thus, the IMSI is initially sent unprotected over the link between the UE and the SN in LTE. Also in 3G and 2G the IMSI is initially sent unprotected over the link between the UE and the SN. But in 5G the USIM may also store the public key *pk* of the HN; and then the UE, before identifying itself, can send encrypted (by the key *pk*) message to the HN.

When a UE wants to connect to a SN, the SN wants to know the identity of the user so that the user can be billed. The SN sends an IMSI inquiry to the UE. Since the link between the UE and the LTE (also 3G and 2G) SN is initially unprotected, the UE has to respond with its IMSI in cleartext. A passive IMSI catcher listens to the radio channel and waits for IMSIs sent in cleartext. An active IMSI catcher impersonates a legitimate SN and makes an IMSI inquiry. The UE has no way to distinguish an active IMSI catcher from a legitimate SN before authenticating the SN. Hence, the UE invariably sends the IMSI to the attacker in cleartext.

The identification is followed by mutual authentication based on challenge and response. We will now outline the basic principle of the authentication protocol. The SN provides the identity of the user to the HN. The HN (i) prepares a random challenge; (ii) computes the expected response to the challenge, an authentication token, and an anchor key; (iii) may compute some other required information. The response, authentication token and anchor key are computed using the master key K. An authentication token includes information that protects the integrity of the challenge and defeats replay attack. The challenge, response, authentication token, anchor key and also some other relevant information are collectively known as an authentication vector (AV). The HN sends the AV to the SN.

The SN sends the challenge and the authentication token to the UE. The UE verifies the integrity and freshness of the challenge using the master key K. If the verification result is positive, the UE computes the expected response to the challenge using the master key Kand sends the response to the SN. If the expected response that the SN received from the HN matches with the response that the UE sent, the authentication is successful.

The authentication and key agreement protocols of LTE and 5G networks – known as LTE AKA [16] and 5G AKA [6] – are based on the above principle. In LTE AKA, if the authentication is successful, and if the UE that participated in the AKA is attaching with the SN for the first time, then the SN sends a location update (LU) message to the HN [12, 17]. Once a UE is authenticated, the SN gives it a temporary identity – known as globally unique temporary UE identity (GUTI) – with confidentiality and integrity protection over a secure channel that has been established based on the authentication. Since then, the UE uses GUTI to identify itself to the SN. The use of GUTI defeats the passive IMSI catchers. However, the use of GUTI does not defeat active IMSI catchers. This is because the UE or the SN may have lost the GUTI, or the UE may visit a new SN. In either case, the SN would make an IMSI inquiry towards the UE. The UE would have to respond with IMSI in cleartext, otherwise it would be locked out of the network.

Once a UE is attached to an SN, the UE may go to idle mode. When there is downlink data available for an idle UE, the SN would page the UE to wake it up, using the IMSI or the GUTI (of that UE) in the paging message. It is worth mentioning that a new authentication may be in initiated by the SN after UE responds to a paging message.

3 Related Work

The attack on user identity privacy by active IMSI catchers has been known since GSM, and numerous defence mechanisms against this attack have been proposed in the literature. We identified two major trends in these mechanisms: the first is based on pseudonyms and the second on public key cryptography. Below, we list works that we consider to be the most relevant to this paper.

In Herzberg et al. [18], one-time pseudonyms are created by probabilistic encryption of the user's identity using a key that is known only by the HN. The HN sends a set of pseudonyms to the UE when the UE and the HN have a secure communication channel.

Publications proposing cellular network pseudonyms in the same format as IMSI, but with randomized, frequently changing MSIN, include [7–10]. In Broek et al. [7] and Khan and Mitchell [8] the pseudonym's update is embedded in a random challenge, RAND, of AKA. Khan and Mitchell [11] identified a weakness in solutions proposing cellular network pseudonyms in the same format as IMSI, which could be exploited to desynchronize pseudonyms in the UE and the HN. They also proposed a fix. Khan et al. [12] found a weakness in [11] and proposed a solution. We will refer to the solution in [12] as the KJGN scheme.

Asokan [19] described how public key encryption can be used to achieve identity privacy in mobile environments. In this solution, only the HN has a public/private key pair and the UE is provisioned with the public key of the HN. The UE encrypts identity information using public key before sending it to the HN. Køien [20] suggests using identity based encryption (IBE) to defeat IMSI catchers in LTE. Khan and Niemi [21] propose the use of IBE to defeat IMSI catchers in 5G networks.

5G is the first generation of mobile networks that includes protection against active IMSI-catchers. 3GPP has decided that users' identities will be protected in 5G by including public key encryption [6]. The idea is similar to Asokan [19]. In addition, the possibility to include IMSI in the paging message has been removed. We will now outline the working of 5G protection mechanism [6].

The UE conceals the 5G user's long-term identities with Elliptic Curve Integrated Encryption Scheme (ECIES) [22, 23] before sending them to the SN. ECIES is a hybrid encryption scheme that combines an elliptic curve based public key cryptography with secret key cryptography; it is a semantically secure probabilistic encryption scheme ensuring that successive encryptions of the same plaintext with the same public key result in different ciphertexts with very high probability. Specifically, [6] includes two ECIES profiles, both for the approximately 128-bit security level. Both profiles use AES-128 [24] in CTR mode [25] for confidentiality and HMAC-SHA-256 [26,27] for authenticity in the secret key cryptography part but use either Curve25519 [28, 29] or secp256r1 [23] elliptic curves for the public key cryptography part.

A UE that is provisioned with the HN's public key pk,[3] uses it to construct SUCI by computing $\mathcal{E}_{pk}$(MSIN)—the encryption of MSIN with the HN's public key pk—and concatenating it with certain cleartext information. As defined in [6], this cleartext information includes HNID, the home network identifier enabling successful routing to the HN, SUPIPSI for defining the scheme (i.e., either a null scheme or the ECIES profile), and HNPKI for denoting which public key was used in encryption. ECIES guarantees that MSIN can be decrypted from SUCI only by the HN who holds the private key corresponding to pk. Hence, the users' identities are protected.

Please note that public-key based solutions, where the IMSI is delivered to the network encrypted by the public key of the network, are conceptually simpler than pseudonym-based solutions, because the latter need a mechanism for changing pseudonyms while keeping the set of pseudonyms in the UE and the

[3] The standard [6] does not require the HN to provision pk into every UE. If HN has not provisioned its pk into a UE, then that UE will not conceal its long-term identity with this mechanism.

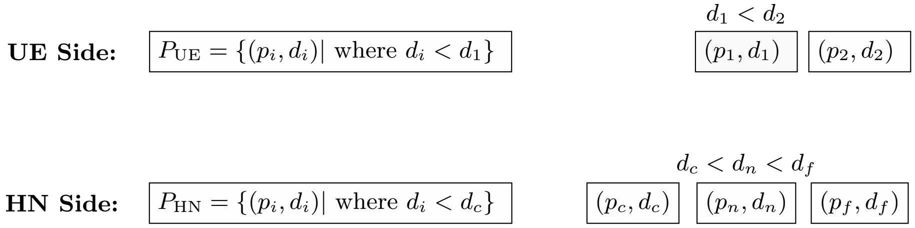
$d_1 < d_2$
UE Side:
$P_{\mathrm{UE}} = \{(p_i, d_i)|\text{ where } d_i < d_1\}$
(p_1, d_1)
(p_2, d_2)
$d_c < d_n < d_f$
HN Side:
$P_{\mathrm{HN}} = \{(p_i, d_i)|\text{ where } d_i < d_c\}$
(p_c, d_c)
(p_n, d_n)
(p_f, d_f)

Fig. 2: Pseudonyms in UE and HN

HN synchronized. On the other hand, the impact of pseudonyms in IMSI format on legacy network nodes between UE and HN is smaller than of IMSI encrypted by a public key of the network, because the format of IMSI encrypted by a public key of the network is quite different from plaintext IMSI.

4 Our Solution

The 5G standard protects long-term user identity against active IMSI catchers by using SUCI (generated by the public key of the HN) [6], as outlined in Section 3. Despite this protection in 5G, a fake LTE SN can still mount a downgrade attack on the identity privacy of a 5G UE.

We mitigate the downgrade attack as follows: instead of IMSI, a 5G UE in LTE network uses pseudonyms that have the same format as IMSI. In addition, when a 5G UE runs 5G AKA, it also synchronizes its LTE pseudonyms with the HN. In our solution:

1. a 5G UE uses pseudonyms to connect with LTE SNs and Release-15 SUCI to connect with 5G SNs;
2. only the HN allocates and releases pseudonyms of mobile users; initially the HN allocates two pseudonyms per 5G user and provisions them into those users' USIMs;
3. a 5G UE gets new pseudonyms by participating in authentication protocols: LTE AKA or 5G AKA; the two latest pseudonyms received by the UE during successful AKA are denoted by p_1 and p_2;
4. in order to support simultaneous connections with several SNs, our solution:
   (i) uses subscriber-specific counter d of pseudonyms maintained by the HN;
   (ii) keeps track of in-use pseudonyms in 5G UE and HN, using sets P_{UE} and P_{HN}, respectively; the elements of these sets are pairs (p_i, d_i) of pseudonyms and their respective counters (see Figure 2);
   (iii) piggybacks information about pseudonyms in 5G UE within the Release-15 SUCI.

The 5G UE will use only p_1 or p_2 when replying to IMSI inquiry from an LTE SN. The set P_{UE} contains pseudonyms that 5G UE received before p_1 and p_2. The UE deletes pseudonyms from P_{UE} based on policy provided by the HN that could include, e.g., pseudonyms' lifetime or the maximum size of P_{UE}. The P_{HN} contains pseudonyms that the HN thinks are in P_{UE}, and it deletes pseudonyms from P_{HN} according to another policy. One objective of these policies is that the HN should not delete a pseudonym which the UE has not deleted yet. Thus, the pseudonyms in the UE constitute a subset of that UE's pseudonyms in the HN. In short, the UE informs the HN about its oldest pseudonym when it connects with a 5G SN using SUCI, and then the HN is able to reduce the set P_{HN}. Please note that as long as the UE is connecting to LTE SN only, the size of P_{HN} grows. It will be explained later how to avoid that P_{HN} grows too much.

Our solution does not modify the structure and/or length of any existing message. Also, it does not introduce any new messages on top of what 3GPP has standardized; it only introduces changes in the 5G UE and its HN. However, to enable lawful interception in the SN, we would need some changes in the LTE SNs too. This issue is discussed in Section 5.2.

As standardized, a 5G USIM comes with an IMSI, a master key K and the HN's public key pk embedded in it. A 5G USIM in our solution also has to include two pseudonyms p_1, p_2 and a key κ, shared with HN, for decrypting the pseudonyms. Similarly, along with the user's IMSI, and master key K, the HN in our solution has to store additional information: the shared key κ for encrypting pseudonyms and three pseudonyms $p_c, p_n,$ and p_f (where the subscripts stand for "current," "next" and "future"). Ideally $p_1 = p_c$ and $p_2 = p_n$.

When a pseudonym p is allocated to a subscriber, it is associated with a subscriber-specific counter d. We require that d is a strictly monotonically increasing integer variable that increases each time the HN allocates a new pseudonym to the subscriber.

4.1 LTE AKA Based Solution

This solution is shown in Figure 3. It is built on top of the KJGN scheme, which was built on top of LTE AKA. The differences to LTE AKA are indicated by darker font in the figure. The additions on top of KJGN scheme include the use of the counter d_i, which is associated with pseudonym p_i, and the sets $P_{\mathrm{UE}}, P_{\mathrm{HN}}$. One major modification is in the condition on which a UE or the HN forget pseudonyms. Supplements for lawful interception are not shown in Figure 3; they will be discussed separately in Section 5.2.

Description.

(1) An LTE SN inquires the UE about the IMSI.
(2) The UE chooses one of the pseudonyms p_1, p_2 and assigns it to q.
(3) The UE sends q to the SN.
(4) The SN sends an AV request for the pseudonym q to the HN. It is worth mentioning here that in most of the times the user identifies itself with GUTI. Sometimes the user would implicitly identify itself by responding to a paging message. In either case, if the SN wants to perform an LTE AKA, the SN

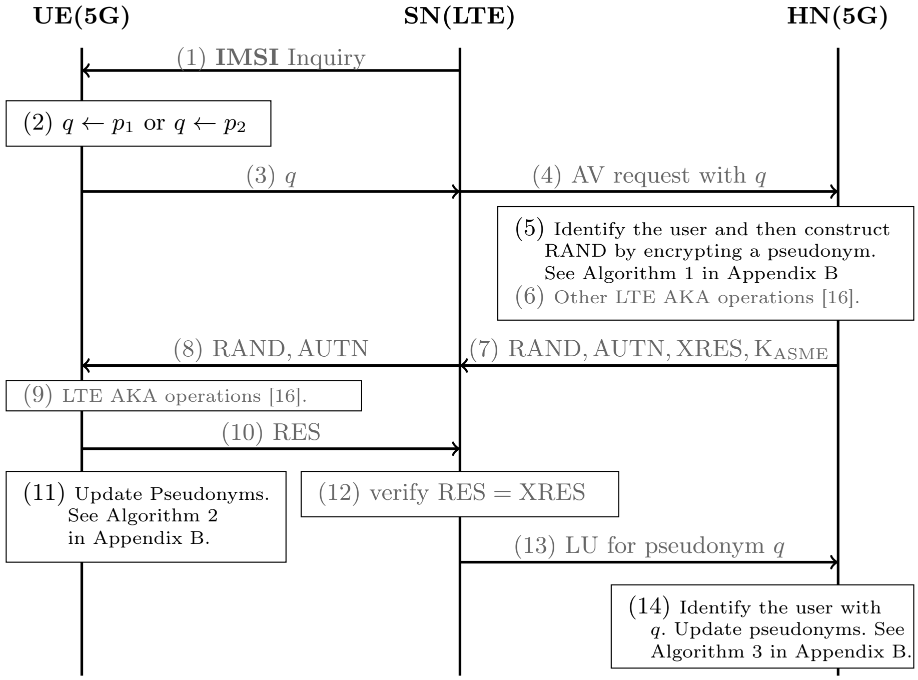
UE(5G)
SN(LTE)
HN(5G)
(1) IMSI Inquiry
(2) $q \leftarrow p_1$ or $q \leftarrow p_2$
(3) q
(4) AV request with q
(5) Identify the user and then construct RAND by encrypting a pseudonym. See Algorithm 1 in Appendix B
(6) Other LTE AKA operations [16].
(8) RAND, AUTN
(7) RAND, AUTN, XRES, K_{ASME}
(9) LTE AKA operations [16].
(10) RES
(11) Update Pseudonyms. See Algorithm 2 in Appendix B.
(12) verify RES = XRES
(13) LU for pseudonym q
(14) Identify the user with q. Update pseudonyms. See Algorithm 3 in Appendix B.

Fig. 3: Solution, when the SN is from an LTE network.

requests an AV to the HN for the pseudonym/IMSI that was associated with the GUTI or was used in the paging message.

(5) The HN checks if the pseudonym q is in use for any subscriber. If yes, the HN starts to prepare the AV. It first constructs the random challenge RAND. A new pseudonym is embedded (encrypted with key κ) in the RAND. Detail of how RAND is constructed is presented in Algorithm 1 in Appendix B. We also describe it in the following.

- The 128-bit long random challenge RAND is created by encrypting (using key κ that can be generated from the master key K) the pseudonym p_f, its counter d_f, an error correction flag (ECF) and a randomly chosen l-bit long *salt*. If the pseudonym p_f is null, a new m-bit long p_f is chosen randomly from the pool of unused pseudonyms. Then d_f is set to the current value of the counter CTR, which is a strictly monotonically increasing counter maintained by the HN. It increases each time the HN generates a new pseudonym. The flag ECF is by default set to 0 but a 5G HN may set it to some other values to notify the UE about an error in the UE's pseudonym state.
- The value of l is equal to $(128 - \text{length}(d_f) - \text{length}(ECF) - m)$. The value of m depends on how many digits of the IMSI are randomized. Since the number of randomized digits can be at most 10, $m \leq 34$. The

length of d_f and ECF depends on implementation; length(d_f) ≤ 24 and length(ECF) ≤ 2 bits should be enough. This implies $l \geq 68$.

(6) The HN computes other parts of authentication vector AV (in addition to the RAND), e.g., the expected response XRES to the challenge RAND, anchoring key $\mathrm{K}_{\mathrm{ASME}}$, and authentication token AUTN [12,16].

(7) The HN sends RAND, AUTN, XRES and $\mathrm{K}_{\mathrm{ASME}}$ to the SN.

(8) The SN forwards RAND and AUTN to the UE.

(9) The UE performs LTE AKA related operations, e.g., verifying MAC in AUTN, computing response RES* [12,16].

(10) If everything is fine in Step (9), the UE sends RES to the SN.

(11) The UE decrypts RAND to extract embedded pseudonym p, and the counter d; and updates the pseudonyms p_1, p_2 if p is new. These operations are presented in Algorithm 2 in Appendix B, and also described in the following.
- UE decrypts RAND using key κ and gets p, d, ECF and *salt* (Line 1).
- In an LTE AKA, the *ECF* bit is always set to 0.
- If the pseudonym p is a new pseudonym i.e., $d > d_2$, then the UE inserts (p_1, d_1) into P_{UE} and sets $(p_1, d_1), (p_2, d_2) \leftarrow (p_2, d_2), (p, d)$. See lines through 7-10.
- If $d \leq d_2$, then p is considered as an old pseudonym and the UE does not update pseudonyms.
  - If somehow the value of d_2 (in the UE) gets corrupted and becomes larger than d_f (in the HN), the UE would never be able to accept new pseudonyms anymore just by running LTE AKA. But even then the UE can still obtain a new pseudonym by running a 5G AKA.

(12) The SN compares RES and XRES; the SN stops if RES $\neq$ XRES.

(13) The SN sends an LU message to the HN for the pseudonym q.

(14) The HN searches for a user s.t., $q \in \{p_n, p_f\}$. If found, and p_f is not null, then the HN inserts (p_c, d_c) into P_{HN} and sets $(p_c, d_c), (p_n, d_n), (p_f, d_f) \leftarrow (p_n, d_n), (p_f, d_f), (NULL, NULL)$. See Algorithm 3 in Appendix B.

4.2 5G AKA Based Solution

This solution is used for: (i) delivering a new pseudonym to a 5G UE using 5G AKA; (ii) notifying the HN about pseudonyms that the UE is not using anymore; so that those pseudonyms can be reused in HN; and (iii) re-synchronization of pseudonym states in the (rather unlikely) erroneous situation where d_2 becomes greater than d_f. It is required to deliver new pseudonyms to a 5G UE even when the UE has not used the existing pseudonyms to connect with a legitimate LTE SN. This is because, the UE may have used those pseudonyms with an active IMSI catcher. If a 5G UE always connects with a 5G SN, and does not get new pseudonyms by participating in 5G AKA, then the 5G UE will have the same pseudonym for long time. So, if an active IMSI catcher makes many IMSI inquiries over this long time, then the UE would respond to each of those IMSI inquiries with the same pseudonym. Thus, an active IMSI catcher would be able to track and monitor the user with this long-lived pseudonym.

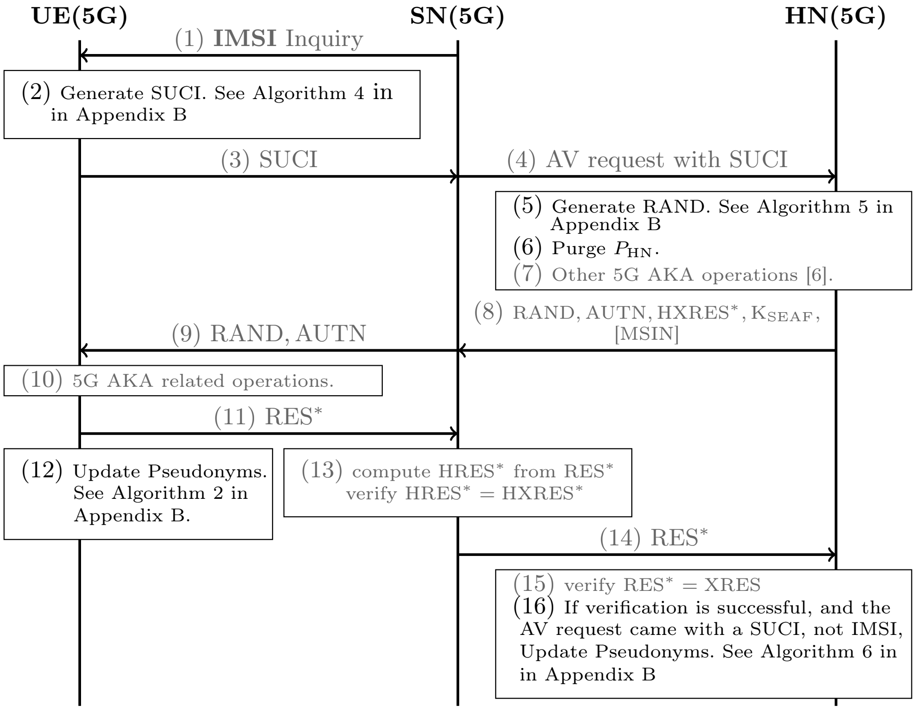
UE(5G)
SN(5G)
HN(5G)
(1) IMSI Inquiry
(2) Generate SUCI. See Algorithm 4 in in Appendix B
(3) SUCI
(4) AV request with SUCI
(5) Generate RAND. See Algorithm 5 in Appendix B
(6) Purge P_{HN}.
(7) Other 5G AKA operations [6].
(8) RAND, AUTN, HXRES*, K_{SEAF}, [MSIN]
(9) RAND, AUTN
(10) 5G AKA related operations.
(11) RES*
(12) Update Pseudonyms. See Algorithm 2 in Appendix B.
(13) compute HRES* from RES* verify HRES* = HXRES*
(14) RES*
(15) verify RES* = XRES
(16) If verification is successful, and the AV request came with a SUCI, not IMSI, Update Pseudonyms. See Algorithm 6 in in Appendix B

Fig. 4: Solution when the SN is from a 5G network.

This solution is built on the 5G AKA protocol of Release 15 [6]; with changes only in the 5G UE and the HN. Thus, the solution is transparent to the 5G SNs of release 15. The solution requires the HN and the USIM to contain all the information the LTE AKA based solution (see Section 4.1) requires. Moreover, it requires the HN to have a public/private key pair pk, sk and the USIM to be provisioned with the HN's public key pk. The solution is presented in Figure 4. The changes in 5G AKA are marked by darker texts.

Description.

(1) A 5G SN inquires the UE about the IMSI.

(2) The UE generates a SUCI. Construction of SUCI by encrypting MSIN using the HN's public key pk is discussed in Section 3 and in [6]. We bring changes in the plaintext that is encrypted into SUCI. Along with MSIN, we also encrypt two counters: δ_{min} and δ_{max}. Here, δ_{min} is the smallest counter of all the pseudonyms in the UE. Thus, the HN can know, which pseudonyms the UE is not using anymore; consequently they can be allocated to other UEs. The value of the counter δ_{max} is always set to d_2. Our construction of SUCI is presented in Appendix B, Algorithm 4, and also described in the following.

- UE computes MAC T of message $\mathit{MSIN}||\delta_{\min}||\delta_{\max}$ with master key K (Line 1).
- UE encrypts $\mathit{MSIN}||\delta_{\min}||\delta_{\max}||T$ with HN's public key pk.
- The ciphertext is concatenated with other information that are in plaintext: HN identifier, the public key identifier of the HN, and the SUPI protection scheme identifier (Line 2). The outcome is returned as SUCI.

(3) The UE sends SUCI.

(4) The SN forwards the SUCI to the HN, requesting an AV. Note that, in most of the times the user identifies itself with GUTI. Sometimes the user would implicitly identify itself by responding to a paging message. In either case, if the SN wants to perform a 5G AKA, the SN requests an AV to the HN with the IMSI that was associated with the GUTI or in the paging.

(5) The HN constructs the RAND by embedding a pseudonym in it. The detail is presented in Algorithm 5 as described in the following.

- The HN extracts MSIN, $\delta_{\min}, \delta_{\max}$ and T from the encrypted part of the SUCI using the private key sk of the HN (Line 1).
- Verifies the MAC T using master key K (Line 2). If the verification is unsuccessful, the algorithm stops. If the verification is successful, the algorithm continues as following.
- Checks if p_f is NULL. If yes, an m-bit long p_f is randomly allocated (from the pool of free pseudonyms) and d_f is set to CTR (Lines 4, 5). CTR is a subscriber-specific counter maintained by the HN. It increases every time the HN generates a new pseudonym.
- Sets ECF to 0. Then checks if $\delta_{\max}$ is greater than d_f. If yes, it sets ECF to 1. See Lines from 7 to 10.
- An l-bit long random salt is chosen (Line 11).
- (p_f, d_f) is set into (p, d) i.e., $(p, d) \leftarrow (p_f, d_f)$. See line 12.
- $(p, d, \mathit{ECF}, \mathit{salt})$ is encrypted with key κ. The ciphertext is RAND. (Line 13). It is worth mentioning that the key κ can be derived from the master key.
- The value of m and l is discussed in Section 4.1.

(6) HN removes pseudonyms from P_{HN} which have counter smaller than $\delta_{\min}$.

(7) HN performs other operations of 5G AKA except constructing RAND [6].

(8) HN sends an AV $(\text{RAND}, \text{AUTN}, \text{HXRES}^*, \text{K}_{\text{SEAF}}, \text{MSIN})$ to the SN.

(9) SN forwards the RAND and AUTN to the UE.

(10) UE performs 5G AKA related operations e.g., verifying AUTN and computing the response RES^* to the challenge [6].

(11) UE sends the response RES^* to he SN.

(12) UE decrypts RAND; extract the embedded pseudonym from the RAND; and updates the pseudonyms in the UE. Algorithm 2 presents the details. This algorithm is partially described in Section 4.1. In 5G AKA the ECF might be set to 1 by the HN. In this case the UE would empty the set P_{UE}, set $(p_1, d_1), (p_2, d_2) \leftarrow (p, d-1), (p, d)$ and terminates the algorithm at this step. This is needed to recover from a very unlikely error situation where d_2 gets corrupted in the UE. We explain the details in the end of Section 5.1.

(13) SN computes HRES^* as a function of RES^*; compares HRES^* with HXRES^*.

(14) If the comparison in last step matches, SN forwards the RES* to the HN.

(15) The HN compares RES* and XRES.

(16) If the comparison in previous step matches, HN checks whether the AV (associated with the current current 5G AKA run) came with a SUCI or an IMSI. If with a SUCI, HN checks if the pseudonym p that was embedded in the RAND is still p_f. If yes, the HN moves (p_c, d_c) to P_{HN} and sets $(p_c, d_c), (p_n, d_n), (p_f, d_f) \leftarrow (p_n, d_n), (p_f, d_f), (NULL, NULL)$. Consequently, the HN would embed a new pseudonym in the RAND while responding to the next AV request. Details are presented in Algorithm 6 in Appendix B. On the other hand, if the UE identified itself with GUTI or responded to a paging message, then the subsequent AV request sent by the SN would be with an IMSI, not with a SUCI. Consequently the pseudonyms would not be updated in HN. This means, in response to the next AV request, the HN will embed the same pseudonym in the RAND.

Step (16) helps the system to avoid generating unnecessary pseudonyms. If a 5G UE attempts to connect with an LTE SN using a pseudonym, but the subsequent LTE AKA fails or no LTE AKA follows (possibly because the SN is an active IMSI catcher), then the next time the UE tries to connect with a 5G SN, it uses SUCI instead of GUTI. In this way the UE can notify the HN that it needs new pseudonym, and it would receive a new pseudonym in the next AKA. Thus, the solution avoids generating unnecessary pseudonyms.

4.3 Allocation of Pseudonyms

In our solution, the pseudonyms are allocated to users by the HN. A pseudonym must not be allocated to two different users at the same time, because such double allocation would hinder correct call routing and billing; e.g,. a user could receive the bill for services used by another user. As new pseudonyms are generated for a user, the older pseudonyms are stored in the sets P_{HN} and P_{UE}. A pseudonym should not be allocated to a new user as long as it is in P_{HN} and P_{UE} of any other user. If pseudonyms are never removed from P_{HN} and P_{UE}, the system will eventually run out of free pseudonyms. Hence, policies are needed for removing pseudonyms from these sets. One objective of these policies is that, a pseudonym should not be deleted from P_{HN} of a user that is not yet deleted from the set P_{UE} of that user.

The HN randomly allocates a new pseudonym for a user from a pool of free pseudonyms, maintained by the HN. A pseudonym p can be in the pool of free pseudonyms only if it is not in the set P_{HN} (or used as p_c, p_n, or p_f) for any user. To keep the pool of free pseudonyms large enough, the HN needs to remove old pseudonyms that are not anymore used by a UE from P_{HN}. Before removing a pseudonym p from P_{HN}, the HN needs to know that the UE is no longer using it and has removed it from the corresponding P_{UE}.

In our solution, a UE removes pseudonyms from P_{UE} according to the policies provisioned in the UE by the HN. The UE notifies the HN about the pseudonyms that the HN can remove from the P_{HN}. The UE sends (encrypted in the SUCI message) the smallest counter value $\delta_{\min}$ of all the pseudonyms available in the UE. The HN then removes (from P_{HN}) all pseudonyms that have smaller counter values than $\delta_{\min}$. The UE sends $\delta_{\min}$ with both integrity and confidentiality protection within the encrypted part of SUCI, as discussed in Section 4.2.

As mentioned earlier, as long as the UE is connecting to LTE SN only, the size of P_{HN} grows. We will explain next how to avoid that P_{HN} grows too much. The HN needs to have a cap for the size of P_{HN}. If the cap is reached, the HN would not generate any future pseudonym p_f for the user — but will always embed the same pseudonym p_n in the RAND. As a consequence, the user does not enjoy full identity privacy before it connects to a 5G SN using SUCI; the SUCI would include $\delta_{\min}$ and the HN would be able to reduce P_{HN}. The details of using this cap are left for further study. The cap should be large enough so that it takes a reasonably long time for a UE to reach the cap; on the other hand the cap should be small enough so that a set of legitimate but malicious UE can not exhaust the pseudonym space by connecting to LTE SNs many times. Another topic for future work is how to maximize the chance that the UE would connect with a 5G SN using SUCI. For example, the UE may connect to 5G in order to reduce the size of P_{UE}, even if the user is not paying for 5G services. Another way to recover is to send a message to the user about the situation; suggesting to connect to a 5G SN to enjoy better privacy. Other, more sophisticated techniques than capping the size of P_{HN} can also be in the scope for further study.

It is important to define when the UE can decide that it no longer uses a pseudonym and it can be removed from P_{UE}. The pseudonyms in P_{UE} are stored because the UE should be able to respond to paging messages sent by the SN. Therefore, if a UE has a pseudonym (and the associated GUTI) that has not been used for a reasonably long time (as defined in the policy) and the UE is currently connected to a different SN, the pseudonym can be removed.

The UE may have an old pseudonym in P_{UE} that is associated with a GUTI and a security context but has no other pseudonyms associated with the same SN and the UE is currently connected to this SN. In such a case, the UE would initiate a new registration procedure with the SN using pseudonym p_1 or p_2. If this registration is successful, the UE can remove the old pseudonym from P_{UE}. The UE may also follow a guideline set by the HN to remove pseudonyms from P_{UE}; e.g., remove pseudonyms that are older than one day.

How Large the Sets P_{UE} and P_{HN} Can be. In the HN, if 50 % of MSIN space is used by normal IMSI and pseudonyms, and the allocation of a new pseudonym in HN involves (i) generating a random MSIN, and (ii) checking if that MSIN value has been already allocated, then on the average it would take two tries (each try consisting of (i) and (ii)) for the HN to find a free MSIN. Because of efficiency reasons, we do not want to exceed this number of tries on the average. It follows that the target number of tries and the current level of MSIN space allocation determine the maximum for the average number of pseudonyms per user. Thus, the policy of handling the size of P_{UE} has to be adjusted so that the average size of P_{UE} would not exceed a certain limit, which in the worst case could be in the order of ten.

We estimate that only a small fraction of MSIN's space, in the order of 1 % or less, is in currently in use in the major mobile networks. The biggest fraction, roughly 3 %, of the MSIN's space seems to be in use in China Mobile networks, which has three MNC codes [30] and 910 million subscribers [31]. If the average size of P_{HN} is ten, then for each subscriber, around 13 (also p_1, p_2 and IMSI) elements from the MSIN space would be allocated on the average. Therefore, the fraction of in use MSIN's space will grow by a factor of 13, e.g., from 3 % to 39 %.

Alternative Allocation Mechanisms. In our solution, pseudonyms are allocated to users in the HN. Another approach would be to generate pseudonyms on the UE side. However, in this other approach, the HN has to be able to map a pseudonym with the correct IMSI of the user. Here we briefly present few alternative options and their downsides.

In this option, the UE may perform format preserving encryption (FPE) of the MSIN with a shared symmetric key and use the ciphertext along with MCC and MNC to construct the pseudonym. There may be only one shared key for the whole network, or separate shared keys for each small group of users. Only one key for the whole network is not secure because an attacker can easily know the key. Indeed, the attacker would only need to be be a valid subscriber of the HN and the ability to read the UICC. On the other hand, separate shared keys for each small group do not result into good privacy. If the size of the group is k, the subscriber can achieve only k-anonymity. This is because the user would need to identify the user's group in plaintext. In the roaming case, it could be even worse than k-anonymity. For instance, it could be the case that only one member of the group is roaming in a certain country at a certain time point.

Another option is that the UE pseudorandomly generates pseudonyms — e.g., by hashing the IMSI and a salt using the shared master key. The UE uses m bits of the hash digest along with the MCC and MNC a to construct the pseudonym. Similarly, the HN would need to compute the hash digest (of m bits) of all the IMSIs (of all the valids subscribers) with the salt using the respective master keys. If the salt is chosen by the UE based on an agreed scheme (e.g., the salt is the current date), then the HN has to compute hash digest for each subscriber according to that scheme (e.g., once in a day) and store in a hash table. However, since m is only around 34 bits, there would have many collisions — mutiple IMSIs will be hashed to the same pseudonym. The HN would consult with the hash table when it receives a pseudonym. Due to many collisions in the hash table, pseudonyms would be ambiguous.

5 Analysis of Our Solution

In our solution, the pseudonyms are delivered to the UE with confidentiality protection using the key κ. Hence, an IMSI catcher cannot know a pseudonym before the UE uses it. This provides unlinkability across the pseudonyms. Once a UE switches to use a new pseudonym, the UE appears as a new (previously unknown) user in the network. Same pseudonym may be transmitted many times in different RAND. However, the challenge still remains fresh and random because of the randomly chosen l-bit long *salt*; the value of l can be as long as 68 bits. Since a user keeps using the same pseudonym until it obtains a new pseudonym, the UE is exposed to an active IMSI catcher for the time. This exposure mostly coincides with the already existing exposure of the GUTI. In the rest of this section we analyze our solution from important aspects of using IMSI-like pseudonyms in mobile networks: (i) synchronization of pseudonyms, (ii) LI and patching, (iii) billing and charging, and (iv) performance overheads.

5.1 Synchronization

What is Desynchronization? We say that a UE is desynchronized with the HN if the pseudonyms p_1 and p_2 of the user in the UE side are no more associated with the same user in the HN side. In other words, a UE is desynchronized with the HN if the following condition holds.

$$(p_1, d_1), (p_2, d_2) \notin \{(p_c, d_c), (p_n, d_n), (p_f, d_f)\} \cup P_{\text{HN}}$$

The UE is not allowed to use any other pseudonyms than p_1, p_2 in response to an IMSI inquiry from an LTE, 3G or GSM SN. Consequently, when the UE responds to an IMSI inquiry with p_1 or p_2, the HN would fail to retrieve the correct IMSI. As a result, the subsequent AKA would fail and the UE will not be able to join the network.

Can Desynchronization Happen? If both the HN and UE function correctly desynchronization can not happen. We will present our argument in this section. In principle, desynchronization may happen if one of the following cases happen:

1. If UE accepts a pseudonym that is not generated in the HN
2. If HN deletes $(p_1, d_1), (p_2, d_2)$ from $\{(p_c, d_c), (p_n, d_n), (p_f, d_f)\} \cup P_{\text{HN}}$
3. If UE accepts a pseudonym that was generated for the user in the HN but the HN has already deleted the pseudonym

First, we discuss Case 1. Since a pseudonym is embedded in the RAND, the integrity of the pseudonym is protected if the integrity of the RAND is protected. The integrity of the RAND with the help of the MAC that is a part of the AUTN. So, the UE would never accept a pseudonym that was not generated in the HN; consequently, Case 1 can never happen.

Second, we discuss Case 2. The HN never deletes $(p_c, d_c), (p_n, d_n), (p_f, d_f)$. The HN may delete a pseudonym (p_i, d_i) from P_{HN} if the HN has decrypted a $\delta_{\min}$ from a valid SUCI such that $d_i < \delta_{\min}$; see Step 6 of 5G AKA based solution in Section 4.2. Since, the UE would never deletes $(p_1, d_1), (p_2, d_2)$, the value of $\delta_{\min}$ would be at most d_1. This implies $\delta_{\min} \leq d_1 < d_2$. Consequently, the HN would not delete $(p_1, d_1), (p_2, d_2)$ from P_{HN}.

Third, we discuss Case 3. As discussed in Case 2, if pseudonym (p_i, d_i) is deleted from P_{HN}, then d_2 would be greater than d_i, because $d_i < \delta_{\min} \leq d_1 < d_2$. Remember that a UE would accept a pseudonym (p_i, d_i) only if $d_i > d_2$. Thus (p_i, d_i) would never be accepted by the UE.

Resynchronization. However, if due to some unlikely errors in UE and HN, desynchronization happens, our solution can bring back a desynchronized user into a synchronized state automatically just by connecting to a 5G SN. This is because a 5G UE does not need a valid pseudonym to participate in a 5G AKA, and the UE would get a valid pseudonym if it can participate in a valid 5G AKA. This is a major advantage because, as discussed in [12], such a desynchronized user would otherwise have to go to a mobile network operator's shop to change the UICC.

Also, if d_2 in the UE gets corrupted and becomes larger than d_f, then the UE would not be able to accept a new pseudonym; see Algorithm 1,2 and 5. If d_2 becomes a large enough value, then the UE might not accept new pseudonyms anymore. However, the 5G UE is able to resynchronize with HN even if such a corruption happens; just by connecting to a 5G SN using SUCI and running a 5G AKA.

Since the UE would embed $\delta_{\max}$ (which is equal to d_2) in the SUCI message, the 5G HN would know (Algorithm 4) if such a corruption has happened, see Algorithm 5. To help the UE fixing the corruption, the 5G HN would set the ECF in the RAND, see Algorithm 5. The UE decrypts the RAND and extracts the pseudonym (p, d). If the ECF is set, the UE would accept the pseudonym even though $d < d_2$, see Algorithm 2. Since, acceptance of such a pseudonym may break the consistency with other pseudonyms in the UE, the UE deletes all other pseudonyms it has. Pseudonyms p_1, p_2 holds the same value as p; d_1 gets the value $d - 1$ and d_2 gets d. Thus the UE goes back to a state similar to the beginning.

5.2 Lawful Interception and Patching

Lawful Interception (LI) involves selectively intercepting communications of individual subscribers by legally authorized agencies. The agency typically provides the long-term identifier of the target UE or service to the network operator, which must have the means to intercept communications of the correct target based on long-term or permanent identifiers associated with that target. There is a requirement that a network operator is able to intercept without the need to rely on another network operator or jurisdiction. In particular, an SN does not need to share LI target identities of roaming UE with an HN and vice versa [32, 33].

In order to support LI in Release 15, where the UE encrypts its identity using the public key of the HN, (1) the HN provides long-term identifier of the UE to the SN during authentication and key agreement procedure; and (2) both UE and SN use long-term identifier of the UE as one of the inputs to the derivation of master session key. The visited network gets the long-term identifier of the UE as a result of (1), while (2) ensures that the UE and visited network can communicate only if both use the same long-term identifier of the UE.

In order to support LI in our pseudonym-based solution, we propose to add (1), and possibly also (2), into the existing legacy core network elements (MME and HSS in LTE; MSC, SGSN, and HLR in 3G/GSM) by software upgrade. The scope of this software upgrade appears to be much smaller, compared to adapting Release 15 solution into the existing legacy networks. This is mainly because adapting Release 15 solution to legacy networks is likely to impact also the radio access networks: the format of encrypted identifier in Release 15 is quite different from the cleartext IMSI, while in our solution the pseudonym is in the format of IMSI.

The LI feature (1) can be implemented in our solution as follows: in the AV request, the SN informs the HN that it is patched and when the HN returns the AV, it includes the MSIN as part of the AV.

Adding support for (2) is more complicated, especially in the cases of 3G and GSM. In LTE, the SN extracts the MSIN and uses it as an additional input in deriving the master session key and the UE is also informed that the SN is patched by piggybacking on RAND so that the UE also uses the MSIN in computing the master session key. In 3G and GSM there is no concept of master session key. Therefore, the MSIN has to be used directly in derivation of ciphering and integrity keys.

Please note that this complication appears also in the case where Release 15 public-key based solution would be adapted for legacy networks.

Typically, an organization of mobile network operators, like GSMA (GSM Association), could discuss and agree on a deadline in order to provide enough time for operators to patch their networks.

If some legacy network has not completed the patch by the agreed deadline, it is OK for other operators to ignore that failure and start using the pseudonyms-based protection of identity privacy. The LI in the network that is late in patching would not work fully for roaming 5G UEs.

In the case of a public key-based solution, if some mobile operator is late with a patch, then roaming 5G UEs would not be able to join that network.

5.3 Charging and Billing

Mobile users are charged based on CDRs (call detailed records) that are generated in the serving network. A CDR includes IMSI and records chargeable event (like the service used). The home network then adds charges to the bill of subscriber that is associated with the CDR's IMSI.

After a pseudonym-based solution is adopted, a CDR may contain a pseudonym in place of IMSI. (Indeed, if the UE uses pseudonym in IMSI format when it communicates with SN, and the SN does not get the long-term identifier of that UE – for instance, because it was not patched for LI, then the SN has no other choice but to put the UE's pseudonym into the CDR.) For this reason the home network has to consult a log of pseudonym allocations when it maps from

UE identifiers in received CDRs to the actual subscribers' data [8,12]. That log includes the following information: (i) the pseudonym, (ii) the IMSI of the subscriber whom the pseudonym is given to, (iii) the time when HN allocated the pseudonym to the subscriber, (iv) the time when the subscriber started using the pseudonym (i.e., the time of the first successful AKA run of that subscriber with the pseudonym), (v) the time when the UE notified that it is no longer using the pseudonym, and (vi) list of SNs that the user has attached with using this pseudonym. The home network has to maintain the log at least until the billing is settled, and possibly for longer time, to comply with the local authority's guidelines.

In order to keep the billing accurate, a UE that stops using pseudonym p and deletes it, must also stop using the GUTI and the security context associated with p. Let us illustrate what may happen otherwise. A UE1 that is visiting SN1, removes a pseudonym p from its P_{UE} but keeps the GUTI and the security context. The UE1 also informs the HN (via 5G SUCI) that it has removed the pseudonym p. Therefore, the HN sends the pseudonym p to the pool of free pseudonyms; subsequently, p is allocated to another subscriber and delivered to UE2 in SN2.

If UE1 continues to identify itself in SN1 by the GUTI associated with p as it consumes services, then SN1 will generate a CDR that contains the pseudonym p, which is at that time is allocated to UE2. As a result, UE2 will be charged for the service consumed by UE1 in SN1. This may continue until there is a new AKA run between UE1 and SN1 or until the GUTI expires. However, the HN has enough information in the pseudonym allocation log to resolve the correct UE who created the CDR.

5.4 Performance Overheads

In our solution the pseudonyms' delivery protocol between UE and HN is piggybacked on existing AKA messages. The structure of AKA messages remains the same, but parts of those messages are constructed and interpreted differently. We list here the additional tasks that have to be done by HN and UE.

Most of those tasks have to be done in the HN: First, the HN has to store a set of in-use pseudonyms per subscriber, maintain a pool of free pseudonyms, and be able to allocate pseudonyms from this pool uniformly at random.

Second, as mentioned in Section 5.3, the HN needs to log pseudonym assignments for charging purposes. Keeping this log adds to the storage overhead in the HN and computational overhead in the charging system.

Third, the HN has to generate a random l-bit *salt* and encrypt the pseudonym and *salt* using the symmetric key κ. The 5G HN has also to verify the MAC of the message that is encrypted into the SUCI. But the overheads due to these symmetric-key cryptographic operations are negligible.

Fourth, there is one-time, initial provisioning effort into the USIM of: (i) the initial pseudonyms, and (ii) a policy for forgetting pseudonyms.

On the UE side, one extra decryption is needed to extract the pseudonym embedded in RAND. This symmetric-key decryption has negligible overhead. In addition, UE has to maintain a set of pseudonyms based on the policy provisioned by the HN.

6 Conclusion

3GPP Release 15, the first release of 5G system, includes protection of long-term identity of mobile users against active IMSI catching. But in a typical case where 5G UE also supports LTE, it is still vulnerable to LTE IMSI catchers. This threat can be mitigated by adopting pseudonyms in the format of IMSI in LTE.

We propose a solution where the Release 15 mechanism for protecting user identity privacy is used for synchronizing the LTE pseudonyms in the format of IMSI between the UE and HN, thus making the LTE pseudonyms more robust. Our solution can automatically bring a desynchronized user back to synchronized state just by connecting to a 5G network. For LI purpose, the required patching effort in the legacy SNs is reasonable. All in all, pseudonym-based solution is a potential candidate for confounding IMSI catchers in legacy networks.

Questions for future study include the following: (i) Is it a good idea to encrypt the pseudonyms always with the same key? (ii) What data structure can be used to maintain the pool of free pseudonyms? (iii) What part of the solution has to be implemented in the USIM and what part in the ME? Capping the size of the set P_{HN}, when the UE is connecting only with LTE SNs, is also a subject of further study. The resilience of pseudonym-based solution to computational errors in HN, SN, or UE would also be an important area of investigation.

References

1. Soltani, A., Timberg, C.: Tech firm tries to pull back curtain on surveillance efforts in Washington. https://www.washingtonpost.com/world/national-security/researchers-try-to-pull-back-curtain-on-surveillance-efforts-in-washington/2014/09/17/f8c1f590-3e81-11e4-b03f-de718edeb92f_story.html?utm_term=.96e31aa4440b (September 2014) [Online; Was available until 14-July-2017].
2. Intelligence, P.E.: 3G UMTS IMSI Catcher. http://www.pki-electronic.com/products/interception-and-monitoring-systems/3g-umts-imsi-catcher/ [It was available online at least until 6-July-2018].
3. Dabrowski, A., Pianta, N., Klepp, T., Mulazzani, M., Weippl, E.: Imsi-catch me if you can: Imsi-catcher-catchers. In: Proceedings of the 30th Annual Computer Security Applications Conference. ACSAC '14, New York, NY, USA, ACM (2014) 246–255
4. Ney, P., Smith, J., Gabriel, C., Tadayoshi, K.: SeaGlass: Enabling City-Wide IMSI-Catcher Detection. In: Proceedings on Privacy Enhancing Technologies. PoPETs (2017)
5. 3GPP: TS 22.261 Service requirements for next generation new services and markets. https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3107 (2018)

6. 3GPP: TS 33.501 Security architecture and procedures for 5G System. https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3169 (March 2018)
7. Van den Broek, F., Verdult, R., de Ruiter, J.: Defeating IMSI Catchers. In: Proceedings of the 22nd ACM SIGSAC Conference on Computer and Communications Security. CCS '15, ACM (2015)
8. Khan, M.S.A., Mitchell, C.J.: Improving Air Interface User Privacy in Mobile Telephony. In: Second International Conference, SSR 2015, Proceedings, Springer International Publishing (2015)
9. Norrman, K., Näslund, M., Dubrova, E.: Protecting IMSI and User Privacy in 5G Networks. In: Proceedings of the 9th EAI International Conference on Mobile Multimedia Communications. MobiMedia'16, ICST (2016)
10. Ginzboorg, P., Niemi, V.: Privacy of the Long-term Identities in Cellular Networks. In: Proceedings of the 9th EAI International Conference on Mobile Multimedia Communications. MobiMedia '16, ICST (2016)
11. Khan, M., Mitchell, C.: Trashing IMSI Catchers in Mobile Networks. In: Proceedings of the 10th ACM Conference on Security and Privacy in Wireless and Mobile Networks (WiSec 2017), Boston, USA, July 18-20, 2017, United States, Association for Computing Machinery (ACM) (May 2017)
12. Khan, M., Järvinen, K., Ginzboorg, P., Niemi, V.: On De-synchronization of User Pseudonyms in Mobile Networks. In: Information Systems Security, Springer International Publishing (2017) 347–366
13. 3GPP: TS 23.003, 15.4.0 Numbering, addressing and identification. https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=729 (June 2018)
14. 3GPP: TS 23.501 System Architecture for the 5G System. https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3144 (2018)
15. Khan, M., Ginzboorg, P., Niemi, V.: IMSI-based Routing and Identity Privacy in 5G. In: Proceedings of the 22st Conference of Open Innovations Association FRUCT, Jyvaskyla, Finland (May 2018)
16. 3GPP: TS 33.401 V15.0.0 Security architecture (Release 15). https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=2296 (June 2017)
17. 3GPP: TS 23.012 V14.0.0 Location management procedures (Release 14). https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=735 (March 2017)
18. Herzberg, A., Krawczyk, H., Tsudik, G.: On travelling incognito. In: 1994 First Workshop on Mobile Computing Systems and Applications, IEEE Xplore (1994)
19. Asokan, N.: Anonymity in a Mobile Computing Environment. In: 1994 First Workshop on Mobile Computing Systems and Applications, Santa Cruz, California, USA, IEEE
20. Køien, G.M.: Privacy enhanced mutual authentication in LTE. In: 2013 IEEE 9th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob). (Oct 2013) 614–621
21. Khan, M., Niemi, V.: Concealing IMSI in 5G Network Using Identity Based Encryption. In: Network and System Security, Springer International Publishing (2017) 544–554
22. Certicom Research: SEC 1: Elliptic Curve Cryptography. Standards for Efficient Cryptography, Ver. 2.0, http://www.secg.org/sec1-v2.pdf (May 2009)

23. Certicom Research: SEC 2: Recommended Elliptic Curve Domain Parameters. Standards for Efficient Cryptography, Ver. 2.0, `http://www.secg.org/sec2-v2.pdf` (January 2010)
24. NIST: Advanced Encryption Standard (AES). FIPS PUB 197 (2001)
25. NIST: Recommendation for Block Cipher Modes of Operation: Methods and Techniques. SP 800-38A (2001)
26. Krawczyk, H., Bellare, M., Canetti, R.: HMAC: Keyed-Hashing for Message Authentication. RFC 2104 (February 1997)
27. NIST: Secure hash standard (SHS). FIPS PUB 180-4 (2012)
28. Bernstein, D.J.: Curve25519: new Diffie–Hellman speed records. In: Public Key Cryptography (PKC'06). Volume 3958 of Lecture Notes in Computer Science., Springer (2006) 207–228
29. Langley, A., Hamburg, M., Turner, S.: Elliptic Curves for Security. RFC 7748 (January 2016)
30. interactive digital media GmbH: Mobile Country Codes (MCC) and Mobile Network Codes (MNC). `http://www.mcc-mnc.com/`
31. Wikipedia: List of mobile network operators. `https://en.wikipedia.org/wiki/List_of_mobile_network_operators`
32. 3GPP: TS 33.106: 3G security; Lawful interception requirements. `https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=2265` (2018)
33. 3GPP: TS 33.126 V15.0.0 Lawful Interception requirements. `https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3181` (June 2018)

Appendix A Summary of Notation

AUTN authentication token used to verify the integrity of the RAND; computed by the HN; sent to the SN as part of the AV; used in both LTE AKA and 5G AKA

AV authentication vector; a bunch of information that the HN sends to delegate the authentication to the SN

CDR call detailed record

d_i the counter value (time stamp) associated with pseudonym p_i by the HN

$\delta_{\min}$ the smallest counter of all the pseudonyms available in the UE. The UE embeds it in the SUCI.

$\delta_{\max}$ it is always set to the current value of d_2. The UE embeds it in the SUCI.

ECF error correction flag, piggybacked on the random challenge RAND. the 5G HN sets this flag to inform the UE that the UE's pseudonym state is corrupted

5G AKA an authentication protocol used in 5G

GSM Global System for Mobile Communications

HN Home network; the UE is subscribed with this network

IMSI international mobile subscriber identity

κ a symmetric key shared only between the USIM and the HN

K the symmetric key, also known as master key, shared only by the USIM of a user and the HN

$\mathrm{K}_{\mathrm{ASME}}$ anchoring key – subsequent encryption and integrity protection keys are derive from it; computed by the HN and sent to the SN as part of the AV

l bit length of the *salt* that is used as a part of the plaintext that is encrypted into RAND

LTE AKA an authentication protocol used in LTE network

LU location update; when a user attaches to an LTE SN for the first time, after the authentication succeeds, the SN send an LU message to the HN

m number of bits used to generate the randomized decimal digits to construct pseudonyms; $m \leq 40$

ME mobile equipment, usually a mobile phone

p_c, p_n, p_f the pseudonyms in the HN; p_c for the current pseudonym, p_n for new pseudonym and p_f for the future pseudonym; ideally expectation is $p_1 = p_c, p_2 = p_n$ or $p_1 = p_n, p_2 = p_f$

P_{HN} the set of pairs (p_i, d_i) maintained by the HN for a specific UE

p_1, p_2 the pseudonyms the UE is currently using

P_{UE} the set of pairs (p_i, d_i) maintained by the UE

q the pseudonym that a UE uses to identify itself

RAND is the 128-bit long random challenge; chosen/constructed by HN; sent to SN as part of the AV; SN forwards it to the UE; used in both LTE AKA and 5G AKA

RES response sent by the UE to the random challenge RAND in LTE AKA

SN serving network; the UE connects with this network

SQN the sequence number used by the USIM and the HN, both in LTE AKA and 5G AKA to defeat replay attack.

UE user equipment (USIM + ME)

UICC universal integrated circuit card

USIM universal subscriber identity module

XRES expected response (to RAND) from the UE; computed by the HN; sent to the SN as part of the AV; used in LTE AKA

XSQN the sequence number sent by the HN to the UE

Appendix B Algorithms

Algorithm 1 Construct RAND for LTE AKA in HN

Input: $q, (\text{IMSI}, \kappa, (p_c, d_c), (p_n, d_n), (p_f, d_f))$, CTR; where q is the pseudonym received in the AV request, the vector following q contains the relevant records of the user in the subscription database s.t., $q \in \{\text{IMSI}, p_c, p_n, p_f\}$ **or** $(q, *) \in P_{\text{HN}}$, and CTR is a non decreasing counter that the HN maintains. CTR increases once in a configured time interval.

Output: RAND

```
if $p_f = NULL$ then
    allocate $p_f \in \{0,1\}^m$ randomly from the pool of free pseudonyms
    $d_f \leftarrow CTR$
end if
choose $salt \in \{0,1\}^l$ randomly
$(p, d) \leftarrow (p_f, d_f)$
$ECF \leftarrow 0$                ▷ this flag might be set to 1 by a 5G HN to indicate an error
RAND $\leftarrow E_\kappa(p, d, ECF, salt)$
return RAND
```

Algorithm 2 Update Pseudonyms in UE

Input: RAND, p_1, p_2

Output: updated pseudonym states in the UE

```
extract $p, d, ECF, salt$ by decrypting RAND using key $\kappa$
if $ECF = 1$ then                ▷ it means, the UE has unreasonably large $d_2$
    empty the set $P_{\text{UE}}$
    $(p_1, d_1), (p_2, d_2) \leftarrow (p, d-1), (p, d)$
    return
end if
if $d > d_2$ then
    $P_{\text{UE}} \leftarrow P_{\text{UE}} \cup \{(p_1, d_1)\}$
    $(p_1, d_1), (p_2, d_2) \leftarrow (p_2, d_2), (p, d)$
end if
```

Algorithm 3 Update Pseudonyms in HN after LTE AKA

Input: $q, (P_{\text{HN}}, (p_c, d_c), (p_n, d_n), (p_f, d_f))$; where q is the pseudonym received in the AV request, the vector following q contains the records of the user in the subscription database s.t., $q \in \{p_n, p_f\}$

Output: updated pseudonym states in the HN

```
if $p_f \neq NULL$ then
    $P_{\text{HN}} \leftarrow P_{\text{HN}} \cup \{(p_c, d_c)\}$
    $(p_c, d_c), (p_n, d_n), (p_f, d_f) \leftarrow (p_n, d_n), (p_f, d_f), (NULL, NULL)$
end if
```

Algorithm 4 Generate SUCI

```
Input: $\mathrm{MSIN}, \mathrm{K}, \delta_{\min}, \delta_{\max}, \mathrm{HNID}, pk, \mathrm{HNPKI}, \mathrm{SUPIPSI}$; where $\delta_{\min}$ is the counter of
    the earliest pseudonym in $P_{\mathrm{UE}}$; $\delta_{\max}$ is $d_2$; HNID is the HN ID, usually MCC||MNC;
    $pk$ is the public key of the HN; HNPKI is the public key identifier of the HN;
    SUPIPSI is the SUPI protection scheme identifier
Output: SUCI
$T \leftarrow \mathrm{MAC}_{\mathrm{K}}\left(\mathrm{MSIN}||\delta_{\min}||\delta_{\max}\right)$
$\mathrm{SUCI} = \mathrm{HNID}||\mathrm{HNPKI}||\mathrm{SUPIPSI}||\mathcal{E}_{pk}\left(\mathrm{MSIN}||\delta_{\min}||\delta_{\max}||T\right)$
return SUCI
```

Algorithm 5 Construct RAND for 5G AKA

```
Input: $sk, \mathrm{SUCI}, \kappa, (p_f, d_f), CTR$; where $sk$ is the private key of the HN, SUCI is
    received from the SN in the AV request, $\kappa$ is the subscriber-specific key to encrypt
    pseudonyms, $(p_f, d_f)$ is the subscriber's pseudonym and its counter that would be
    embedded in RAND
Output: RAND
extract $\mathrm{MSIN}, \delta_{\min}, \delta_{\max},$ and $T$ by decrypting SUCI using secret key $sk$
if $T = \mathrm{MAC}_{\mathrm{K}}\left(\mathrm{MSIN}||\delta_{\min}||\delta_{\max}\right)$ then
    if $p_f = NULL$ then
        allocate $p_f \in \{0,1\}^m$ randomly from the pool of free pseudonyms
        $d_f \leftarrow CTR$
    end if
    $ECF \leftarrow 0$
    if $\delta_{\max} > d_f$ then                ▷ It means the UE has unreasonably large $d_2$
        $ECF \leftarrow 1$
    end if
    choose $salt \in \{0,1\}^l$ randomly
    $(p, d) \leftarrow (p_f, d_f)$
    $\mathrm{RAND} \leftarrow E_{\kappa}\left(p, d, ECF, salt\right)$
    return RAND
end if
```

Algorithm 6 Update Pseudonyms in HN after 5G AKA

```
Input: $p, (P_{\mathrm{HN}}, (p_c, d_c), (p_n, d_n), (p_f, d_f))$; where $p$ is the pseudonym that was embed-
    ded in the RAND of the 5G AKA in question, the vector following $p$ contains the
    relevant records of the user participated in the AKA.
Output: updated pseudonyms in the HN
if $p_f = p$ then
    $P_{\mathrm{HN}} \leftarrow P_{\mathrm{HN}} \cup \{(p_c, d_c)\}$
    $(p_c, d_c), (p_n, d_n), (p_f, d_f) \leftarrow (p_n, d_n), (p_f, d_f), (NULL, NULL)$
end if
```