**3GPP TSG-SA3 Meeting #107-e *S3-220911***

**e-meeting, 16 - 20 May 2022**

**Source:**  **Ericsson, Apple, AT&T, Cable Labs, China Southern Power Grid Co, Convida Wireless LLC, Intel, Interdigital, Johns Hopkins University APL, Lenovo, LGE, Mavenir, MITRE, NCSC, Oppo, Phillips, Samsung, Telefonica, US NIST, US NSA, Verizon Wireless, Xiaomi, ZTE**

**Title:** **New key issue on SUPI length disclosed by SUCI**

**Document for: Approval**

**Agenda Item:** **5.6**

# 1 Decision/action requested

***It is proposed to approve the key issue described in this document.***

# 2 References

[1] 3GPP TS 23.003: "Numbering, addressing and identification".

[2] 3GPP TS 33.501: "Security architecture and procedures for 5G System".

[3] Mihir Bellare, A. Desai, E. Jokipii, Phillip Rogaway: "A concrete security treatment of symmetric encryption ".

# 3 Rationale

According to TS 23.003 [1], subscribers’ permanent identifiers, called SUPIs, are allowed to be in Network Access Identifier (NAI) format -- username@realm. It is likely that many networks will have the username part created from real-world names. Therefore, SUPIs can have significantly different lengths.

Using the Elliptic Curve Integrated Encryption Scheme (ECIES), a user device and the home network agree on a shared key by leveraging the public key of the home network. The user device uses the shared key in a symmetric encryption scheme (AES in counter mode) to encrypt SUPIs, into concealed identifiers, called SUCIs [2].

In the symmetric-key setup, security notions like real-or-random, left-or-right, or semantic security are defined in the context where plaintexts have the same lengths [3]. Though AES counter mode is secure according to these notions, direct use of it is not sufficient to serve an intended purpose of SUCIs -- indistinguishability of SUCIs. This is because SUPIs can have different lengths, and in counter mode, the length of the plaintext and the corresponding ciphertext is the same. Therefore, when two SUPIs have different lengths, their ciphertexts are distinguishable from each other. Also, SUCIs that have rare lengths are easily recognisable.

# 4 Detailed proposal

\*\*\* 1st CHANGE \*\*\*

# 2 References

The following documents contain provisions which, through reference in this text, constitute provisions of the present document.

- References are either specific (identified by date of publication, edition number, version number, etc.) or non‑specific.

- For a specific reference, subsequent revisions do not apply.

- For a non-specific reference, the latest version applies. In the case of a reference to a 3GPP document (including a GSM document), a non-specific reference implicitly refers to the latest version of that document *in the same Release as the present document*.

[1] 3GPP TR 21.905: "Vocabulary for 3GPP Specifications".

[XX] 3GPP TS 33.501: "Security architecture and procedures for 5G System".

[XY] 3GPP TS 23.003: "Numbering, addressing and identification".

[XZ] IETF RFC 7542: "The Network Access Identifier".

[XW] John Preuß Mattsson and Prajwol Kumar Nakarmi. 2021. "Nori: Concealing the Concealed Identifier in 5G". In The 16th International Conference on Availability, Reliability and Security (ARES 2021). Association for Computing Machinery, New York, NY, USA, Article 105, 1–7. DOI:https://doi.org/10.1145/3465481.3470076

[X1] Mihir Bellare, A. Desai, E. Jokipii, Phillip Rogaway: "A concrete security treatment of symmetric encryption".

\*\*\* 2nd CHANGE \*\*\*

## 5.X Key issue: SUPI length disclosed by SUCI

### 5.X.1 Key issue details

IMSI catchers have the goal to attack subscribers’ anonymity, and consequently track and monitor the associated UEs. To defeat the IMSI catchers, the 5G standard has introduced the encryption of a subscriber’s permanent identifier called SUPI into a concealed identifier called SUCI using the public key of the home network [XX]. However, the existing mechanism to compute SUCIs from NAI format SUPIs, when having variable lengths, does not provide indistinguishability of SUCIs when the corresponding SUPIs have different lengths.

Elliptic Curve Integrated Encryption Scheme (ECIES) is used to encrypt a SUPI into a SUCI. ECIES is a hybrid encryption scheme in which the actual encryption is done by a symmetric encryption scheme using a shared key, which is established using the home network’s public key [XX]. For SUCI, the symmetric encryption scheme used is AES-128 in counter mode. It should be noted that in counter mode, the length of plaintext and the corresponding ciphertext are the same. Therefore, if two SUPIs have different lengths, then their SUCIs would be distinguishable from each other based on the length of SUCIs. It should also be noted that in the symmetric-key setup, security notions like real-or-random, left-or-right, or semantic security are defined in the context where plaintexts have the same lengths [X1]. Therefore, the security of AES-128 in CTR mode, according to these notions, is meaningful only when the plaintexts of interest have the same length.

According to clause 2.2A of TS 23.003 [XY], the 5G standard allows the use of Network Specific Identifiers (NSI) as SUPI. An NSI will take the form of a Network Access Identifier (NAI) as defined in clause 28.7.2 of TS 23.003 [XY]. The NAI for SUPI shall have the form username@realm, which can have variable length username, as specified in clause 2.2 of IETF RFC 7542 [XZ]. It is likely that many networks will have the username part created from real-world names because earlier and current uses of such identifiers, e.g., in ISIMs (IP Multimedia Services Identity Module), have been based on real-world names. Researchers in [XW] have analysed the name length data for the whole of Sweden (ten million people) and four regions (Sweden, China, India, and USA) of an internal company, and have found that the length distributions have tails (Figure 5.X.1-1). Therefore, the SUCIs that have lengths far from the mean (either to the left or right) would have low anonymity – in the worst cases, completely distinguishable.



Figure 5.X.1-1: Name-length histogram. fl = first name || last name; fml = first name || maiden name || last name [XW]

### 5.X.2 Security threats

An attacker that is eavesdropping over the air interface can identify and track subscribers with unusual lengths of the username field of variable-length SUPI in NAI format by their SUCI (e.g., relatively short or long SUPIs).

If such an unusual length of the username field is unique to a single subscriber, that SUCI will be uniquely attributed to that subscriber by an adversary.

If there is a group of subscribers with unusual lengths of username fields in their SUPIs, the attacker will be able to infer the membership of those subscribers in such group.

### 5.X.3 Potential security requirements

The 5G system should provide means to enhance the privacy of variable length SUPIs in NAI format when constructing SUCIs using the non-null scheme.

\*\*\* END OF CHANGES \*\*\*