Mobile Tracking in 5G and Beyond Networks: Problems, Challenges, and New Directions

Changlai Du*, Hexuan Yu*, Yang Xiao†, Wenjing Lou*, Chonggang Wang‡, Robert Gazda‡, Y. Thomas Hou*
*Virginia Polytechnic Institute and State University, Blacksburg, VA, USA
†University of Kentucky, Lexington, KY, USA
‡InterDigital, Inc., Conshohocken, PA, USA

***Abstract*—This paper explores mobile tracking as a privacy threat posed by 5G and beyond (5G&B) cellular networks. We reviewed the mobile network operation and protocol design, with a focus on the use of mobile identifiers and localization methods, as they are the key technical enablers of tracking and localization in mobile networks. We note that user location privacy against mobile network operators (MNO) is an extremely challenging problem. While a permanent identifier is the most critical piece of information to pinpoint a user and create linkability among records of a user, using a dynamic identifier presents significant privacy-utility challenges—many mobile applications and national security and emergency response services rely on the knowledge of mobile locations. At the same time, using an obfuscated version of a permanent identifier is only secure against outsider attackers, not the MNOs. We argue that, to protect user privacy against MNOs, a radically new design is necessary and such design must nullify any permanent identifiers to break record linkability and take into consideration of user anonymity, network utility (e.g., incoming call), user accountability, and user privacy awareness and control simultaneously. Lastly, we identify potential new directions as an initial attempt to tackle this challenging problem.**

I. Introduction

Smartphones are ubiquitous in today's society. According to a Cisco report [1], 84% of the world's population owns a smartphone by 2022 and the number of smartphones is estimated to rise to 6.7 billion by 2023 among which 10% will be 5G smartphones. The up-to-date 5G cellular networks would offer almost omnipresent coverage to connect users and things virtually from anywhere at any time. Compared to 4G Long Term Evolution (LTE) networks, 5G cellular networks provide higher speed, lower latency, and greater network capacity [2]. In the meantime, 5G networks are also featured with improved user localization precision due to their small-cell nature and adoption of advanced localization techniques.

User Equipment (UE) tracking is intrinsic to the design of cellular networks for MNOs to provide cellular services. For instance, an incoming voice call to a UE needs to be directed to the UE's connected cell tower first before being broadcast to the UE through a radio channel. This requires UE localization by the MNO at the cell level. On the other hand, MNOs are also incentivized or required by law to share user location data with third parties to provide Location Based Service (LBS). One pioneering LBS is the Enhanced 911 (E-911) emergency service in the US, which was codified into law by the Federal Communications Commission (FCC) in 1996 [3]. Other LBS examples include navigation services, location-based weather services, and location-based advertising applications [4]. The development of 4G and 5G technologies in the past decade also enables an MNO to locate UEs with much finer precision (i.e., compared to cell-level), using advanced physical-layer measurements as well as UE-provided location data [5]–[7].

Though user location data collected by MNOs are essential to cellular network operation and useful to LBS applications, privacy concerns arise with the increasing localization accuracy. The most recent E-911 specification requires that MNOs report the caller's location with 50 meters horizontal accuracy and 3 meters z-axis accuracy in 30 seconds for 80% of the E-911 calls [8]. 5G Standards Release 16 [9] stipulates achieving an accuracy of 3 and 10 meters for indoor and outdoor situations respectively for 95% of the time. With the improved location accuracy, MNOs are able to estimate user locations with address-level precision, which creates a serious privacy problem because user-accessed addresses can reveal a lot of information about the users. For example, by combining a user's home address and publicly available information, one can discover the user's real name and socioeconomic status. MNOs can not only pinpoint real-time locations, but they can also store users' location data history, which can be further compiled to derive the trajectories of individual users. A user's trajectories, combined with other users' trajectories and public databases, can then be used for user profiling or even user behavior prediction.

Subscriber privacy violations and data breaches have been repeatedly reported in recent years. Cellular network operators are reported to routinely sell subscribers' location data to the market. According to a Motherboard report [10], T-Mobile was selling its customers' location data to third parties and as a result, anyone can track every other subscriber's real-time location for just $300. Another example is the lawsuit case *Scott, et al. v. AT&T Inc., et al.* [11], which aims to stop AT&T and two other location aggregators from selling data to other entities—from bounty hunters to car dealerships. The MNOs are the actual sources that create the marketplace where subscribers' real-time location data are traded. The problem is, once the data are sold, they are out of the control of mobile operators, which puts users' privacy at risk no matter if it was MNOs' original intention. In other cases, seemingly innocuous disclosure of location information by mobile apps has led to

2155-6814/22/$31.00 ©2022 IEEE
DOI 10.1109/MASS56207.2022.00067

Fig. 1. Illustration of mobile tracking. Location data samples are collected and stored by MNOs. We define a data sample as a mobility event in the format of a quadruple $< who, where, when, activity >$. MNOs can compile the data samples into user trajectories and further make data inferences using any available tools like artificial intelligence.

serious privacy violations. In a BBC report [12], a fitness app named Strava used a mobile phone's GPS to track a user's exercise activity and visualize the activities of all its users using a heat map according to activity levels. The application was reported to disclose the movements of soldiers at military bases, which might reveal military secrets like possible patrol routes.

In this paper, we study mobile tracking as a systematic privacy threat posed by cellular networks. *Mobile tracking* is the act or process of pinpointing the location or tracking the movement of a mobile device or a person. As illustrated in Fig. 1, we define the location data samples collected and stored by MNOs as *mobility events* in a format of quadruple $< who, where, when, activity >$, among which who is the user's identifier; $where$ and $when$ are location stamp and time stamp respectively; $activity$ is optional, which can be any user operation collected by MNOs such as making a phone call or access to a website. MNOs can compile the data samples into user trajectories and perform trajectory profiling or prediction using any available statistical inference tools like artificial intelligence. Based on this definition, we study mobile tracking by MNOs from two aspects. First, we analyze all the mobile user identifiers in a 5G cellular network to pick out those identifiers that can be used to track users. Second, we inspect the localization techniques used by MNOs to perform mobile tracking in 5G networks. For both aspects, we provide an analysis on how different levels of mobile tracking threats could appear. We then survey the countermeasures proposed in the literature and analyze their strengths and limitations. We posit that the mobile tracking threat posed by the network operators remains an open research area with multiple unsolved challenges, including the trade-off among fine-grained user privacy protection, user accountability, and the handling of incoming calls to anonymous callees. Promising solutions must address all the challenges at the same time with a practical design that adds minimal overhead to the existing mobile infrastructure.

This paper is organized as follows. In Section II, we briefly describe the mobile tracking problem in the bigger picture. In Section III, we analyze mobile tracking by cellular networks

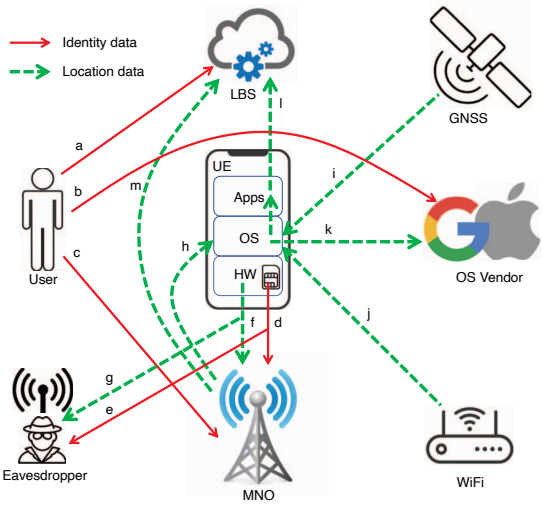

Fig. 2. Mobile tracking diagnosis with information flows. The user and UE can be tracked by multiple entities. The location of the UE can be collected from multiple sources and acquired by multiple entities. This paper highlights mobile tracking threats posed by MNOs.

from both identifier and localization technique perspectives. In Section IV, we review the countermeasures proposed in the literature, followed by a discussion of the design challenges and possible new research directions in Section V. Section VI introduces related works on mobile tracking by other involved entities, and Section VII concludes this paper.

II. Mobile Tracking Context

Using a mobile phone, a mobile user's private information, such as identity and location, is routinely disclosed to multiple service providers. In this section, we first discuss a mobile user's risks of being tracked by various service providers. As shown in Fig. 2, we use a three-layer UE model which includes the physical hardware layer (HW), the operating system layer (OS), and the application layer (Apps) to depict the sharing of user-sensitive information between a mobile user/device and various entities in cellular networks. The lines labeled with alphabets indicate the sharing of information (identity or location) from sources to destinations during normal operations.

A. Identity Sharing

Applications and websites that provide location-based services (LBS) may ask users for registration so that user activities can be tracked and the application service quality can be optimized, as shown in Fig. 2 (line a). If the user prefers to remain anonymous from service providers, he/she can use temporary or fake identities which are discarded after the service is completed. Unlike LBS applications, it is not easy to achieve user anonymity from operating system vendors (line b). This is because OS vendors like Google and Apple require mobile users to have Google Accounts or Apple IDs to use the full features of smartphones. Users are able to hide their real-world identities by registering fake names and fake home addresses. But when a user wants to download apps from Google Play or Apple Store, verified payment methods like

credit cards must be submitted to the OS vendor. Similarly, mobile users have to verify their real-world identities and even credit status to enjoy the post-paid service from mobile operators (line c). Prepaid SIM card registration is not required in the US, but is mandatory in most other countries (157 countries by 2021) [13]. In the US, mobile users can buy "burner phones" from electronic stores with cash to achieve full anonymity from MNOs. A burner phone is usually a prepaid feature phone that is discarded or "burned" after use. Even so, repeated usage of the same phone will increase user profiling risk because MNOs also record the IDs of both the UE and the SIM card (line d). Eavesdroppers like the notorious IMSI catchers [14] may also obtain users' identifiers through the radio channel (line e). But this type of attack has already been mitigated in 5G using an "encrypted" version of the mobile identifier (i.e., SUCI instead of SUPI).

B. Location Sharing

The location information flows are a bit more complicated. MNOs have the power to collect user locations non-intrusively (line f) without the cooperation of UEs. Again, eavesdroppers may perform passive or active attacks to obtain user locations (line g). Location Service in the OS layer of a UE uses cell tower locations (line h), satellite navigation systems (line i), and other wireless signals like Bluetooth and crowd-sourced WiFi hotspots (line j) to estimate UE locations [15]. OS sends user location data to OS vendors (line k) continuously, which makes OS vendors a powerful potential threat to user location privacy. OS vendor threats are not the focus of our study in this paper, but instead, we keep it as one of our future research directions. LBS applications acquire location information through Application Programming Interfaces (APIs) provided by OS (line l). Users can control when and how accurately the applications can get location information. Some applications use functionalities provided by MNOs to collect location information (line m). As we discussed in Section I, this has introduced tremendous risks to subscribers' privacy.

C. The Scope of This Study

This paper places emphasis on the privacy threat of mobile tracking posed by MNOs. This emphasis is drawn based on the following three observations. First, mobile tracking by LBS applications has been well studied and users are typically given an option for their privacy control. We briefly discuss existing research works on this topic in Subsection VI-A. On the contrary, there are few research results published on mobile tracking under the untrusted OS threat model, although many of the privacy-preserving cloud computing works, which assume an honest-but-curious cloud server, are applicable here. For example, trusted execution environment (TEE) technology has been developed and widely deployed in recent years to provide data and program confidentiality and integrity against untrusted servers. We will briefly review this line of research in Subsection VI-B. Second, the user privacy risk posed by MNOs is on the rise as we discussed in Section I. Meanwhile, anti-mobile tracking is a challenging task in cellular networks when we are considering protecting mobile user privacy against the MNOs (i.e., **threat model**: MNO as an honest-but-curious privacy attacker). The **root cause** of the problem lies in the fact that the cellular network architecture is designed to know UE locations in order to provide ubiquitous connectivity. This is the fundamental conflict between the design goals of network utility and user location privacy. There are more conflicts between user identity privacy and cellular network features. Examples include the conflict between anonymity and voice call routing, usage accounting, and misbehaved subscriber accountability. Third, research progress on this important topic has not been encouraging due to the challenges and conflicting design goals. To this end, we try to straighten out the problem, review proposed countermeasures in the literature, and analyze their limitations. Based on the review, we put forward possible research directions.

III. Mobile Tracking Threats by MNOs

The identifier is the most critical piece of information that enables the linkage of a user's multiple mobility events which leads to user tracing and profiling. Meanwhile, the strengthened 5G positioning capability further aggravates the mobile tracking threat. In this section, we first review how the identifiers are designed and used in cellular networks. We then review the positioning techniques of 5G networks, with a focus on network-based positioning methods. Last, we delineate the various levels of mobile tracking threats posed by MNOs.

A. Identifiers in 5G Networks

We summarize the involved identifiers along with their typical communication paths in Fig. 3. Details can be found in 5G specifications [16], [17]. These identifiers are either permanent or temporary, where a permanent identifier keeps a constant value throughout its lifetime, while a temporary one has a value only valid for a period of time.

Subscription Permanent Identifier (SUPI) is the main identifier in 5G networks. SUPI is assigned by the home network (HN), stored in a universal subscriber identity module (USIM) which is issued to the user during service subscription. As indicated by its name, SUPI is a globally unique permanent identifier. HN stores all issued SUPIs and corresponding subscription details. SUPI can either be an International Mobile Subscriber Identifier (IMSI) (E.212 naming plan [18]) as in 4G networks, or a Network Access Identifier (NAI) for non-3GPP radio accesses. SUPI in IMSI format contains Mobile Country Code (MCC) and Mobile Network Code (MNC) which addresses HN to enable roaming scenarios.

SUPI cannot be transmitted in plaintext in the open-air radio channel due to the risk of eavesdropping attacks [14]. Subscription Concealed Identifier (SUCI) is an encrypted version of SUPI generated by UE or USIM using the public key of HN during the UE's initial access registration to or per identity request by the network. SUCI is sent to Access and Mobility Management Function (AMF) in serving network (SN) and then forwarded to Authentication Server Function (AUSF) and then to Unified Data Management (UDM) in

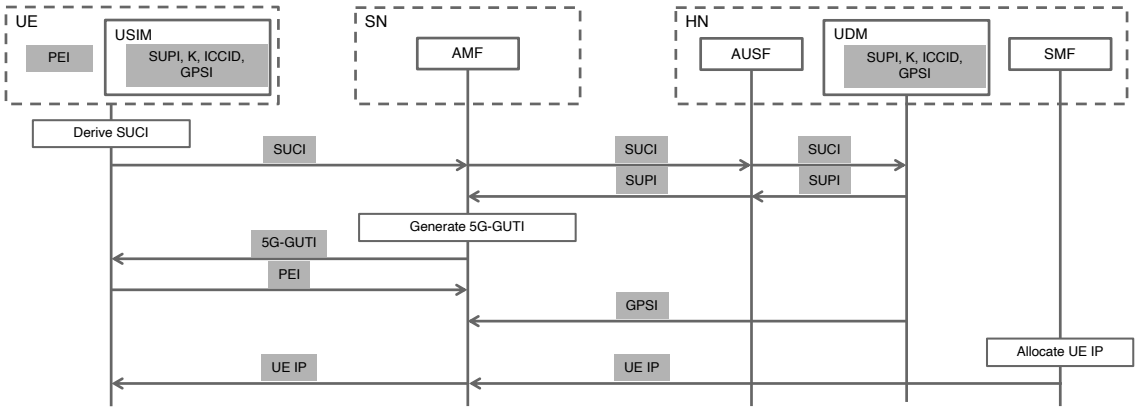

Fig. 3. The identifiers in 5G cellular networks related to user identity privacy. SUPI, K, ICCID, GPSI and PEI are considered permanent identifiers. SUCI and 5G-GUTI are temporary identifiers which are generated by UE/USIM and AMF respectively. UE IP address is assigned by SMF in home network in Home Routing (HR) mode and is temporary is no fixed IP service is purchased.

HN where it can be deciphered into SUPI for authentication. HN will send the SUPI back to AMF if the authentication is successful.

After authentication using SUCI, AMF will allocate a 5G Globally Unique Temporary Identifier (5G-GUTI) to UE which is used for future UE registration. AMF stores the mapping of 5G-GUTI and SUPI. 5G-GUTI is periodically reallocated and the reallocation is mandatory, which protects this temporary identifier from potential attacks [19], [20].

Besides SUPI, a long-term key (i.e., K) is also stored in USIM. K is a root key from which all the user-related keys involved in various 5G communications operations are derived. K can also be used by HN to link a user's multiple accesses as it is a permanent value shared by HN and USIM.

USIM has its permanent hardware identifier known as Integrated Circuit Card Identification Number (ICCID). Similarly, UE has a Permanent Equipment Identifier (PEI), which is equivalent to International Mobile Equipment Identity (IMEI) in 4G LTE networks. PEI is revealed to HN during service subscription. AMF in SN either gets PEI from UE context stored in an old AMF, or it requests PEI from UE during UE registration.

In 5G networks, Session Management Function (SMF) is responsible for allocating an IP address to the UE. SMF is either in SN or HN, depending on the deployed roaming architecture [21]. In either case, SMF passes the UE IP to AMF and then to the UE. In Fig. 3, we follow the Home Routed (HR) architecture (i.e., SMF is in HN) for illustration. The assigned IP address can be dynamic or fixed depending on if a fixed IP service is purchased. In case of a fixed IP is assigned, the IP address is a permanent identifier.

Another identifier that is used both inside and outside of cellular networks is Generic Public Subscription Identifier (GPSI), which is equivalent to Mobile Station Integrated Services Digital Network (MSISDN) (E.164 naming plan [22]) in 4G. MSISDN is the full phone number with the country code included. It can be assigned by HN or the user can transfer a previously owned number to a new HN. The HN maintains a database in UDM that maps GPSI to SUPI for call routing. GPSI is a permanent identifier and is bound to the SIM card as well as the mobile user because a phone number is in a sense also the user's public identifier.

Mobile Tracking Relevance. Based on our previous discussion on identifiers, we make the following observations:

- The permanent identifiers (i.e. SUPI, K, ICCID, GPSI, PEI and Fixed IP) are the coupling points that link multiple mobility events to the same user. Temporary identifiers (i.e. SUCI, 5G-GUTI and dynamic IP) are under control of networks, so they can also be used to link mobility events.
- SUPI is the main identifier. Other identifiers are either correlated to SUPI in cellular network databases, or derived from the value of SUPI. In existing 5G cellular networks, a constant SUPI value is important to the proper operation of the network. For potential privacy measures, SUPI is the most important decoupling point and needs to be nullified to achieve user anonymity. How to nullify SUPI without affecting network functionality is a challenge to anti-mobile tracking designs.
- PEI, which uniquely identifies the UE device, is revealed during SIM registration to HN for UE blacklist checking. How to prevent PEI from being used by the MNO for mobile tracking while retaining the equipment blacklisting feature is a non-trivial task.
- GPSI, which is basically the phone number, can be treated as a part of a user's physical identity because people use their phone numbers heavily as identifiers out of cellular networks and as a public point of contact. Abolishing phone numbers may not be acceptable to users. How to achieve user anonymity while letting the users keep their phone numbers is another design challenge.
- How to remove all the "link points" is still an open question for anti-mobile tracking designs. We will review the research efforts in Section IV-A.

B. 5G Positioning

5G cellular networks collect UE location information at both the cell level and geographic coordinates level [23], [24].

Method	Type	UE-based	UE-asstd	RAN-asstd
NR E-CID	Proximity		●	●
Multi-RTT	Trilateration		●	●
UL-TDoA	Trilateration			●
UL-AoA	Triangulation			●
DL-TDoA	Trilateration	●	●	
DL-AoD	Triangulation	●	●	
A-GNSS	Multilateration	●	●	
WLAN	Hybrid	●	●	
Bluetooth	Hybrid		●	
TBS	Hybrid	●	●	
Sensor	Hybrid	●	●	

Table I. Localization methods in 5G networks. The methods are differentiated by the fundamental positioning techniques. They are also marked if they can be implemented in UE-based, UE-assisted, or 5G-RAN-assisted versions.

Cell-level location information contains the cell ID and Tracking Area (TA) ID of the UE. Coordinates-level positioning information contains UE's geographic coordinates, accuracy estimates, and velocity estimates.

Cell-level location information. When UE is in an active state, its location is known to the network at the cell level. On the other hand, when UE is idle, its location is known to the network at Tracking Area (TA) level. A tracking area comprises a group of cell towers, which can be thought of as the domain to broadcast *paging* messages. Paging is a mechanism to allow the network to send notifications to UE so that UE can wake up to receive the message. The purpose of TA is to reduce the wake-up frequency of UE when the user moves around to save UE energy. In exchange, cellular networks need to broadcast the paging messages on all the towers in the TA.

In 5G networks, one or more TAs can be assigned to a UE as a Registration Area (RA), which serves as a base for the network to locate UE and for UE to update its location to the network. RA can be configured and updated for each UE. This "customized" RA can achieve the best energy efficiency for each UE. A good example is that for a 5G-connected bus, its RA will consist of all the TAs on its route. The bus does not need to update its location to the network as long as it is following the route.

Positioning methods. 5G positioning functionality provides the geographic position and/or velocity of the UE based on signal measurements. We briefly review the standardized 5G positioning methods and highlight the advances in 5G positioning performance. More details can be found in [25].

5G positioning methods are summarized in Table I. The methods can be categorized into UE-based, UE-assisted, and Next Generation Radio Access Network (NG-RAN)-assisted. UE-based means that the UE provides the measurements and also carries out the positioning calculation, whereas UE-assisted and NG-RAN-assisted methods ask UE or NG-RAN to provide the measurements while the 5G core network Location Management Function (LMF) carries out the positioning calculation. For UE-based or UE-assisted methods, the user has the option to disable the positioning functionality. But for 5G-RAN-assisted methods, the UE is non-intrusively positioned, where the user has completely no control.

The basic measurements used for 5G positioning include Radio Signal Strength (RSS), Time of Arrival (ToA), Time Difference of Arrival (TDoA), Angle of Arrival (AoA), and Angle of Departure (AoD). Regardless of the measurements, different positioning techniques can be used to compute the positions, including trilateration, triangulation, proximity, fingerprinting, and hybrid (see [6] for a survey). We list the fundamental positioning techniques in Table I.

In the Cell ID (CID) positioning methods, the position of a UE is estimated using the position of its connected base station gNB. In the Enhanced CID (E-CID) methods, the UE reports additional measurements to improve location estimation. In 5G NR E-CID, the UE is not required to take additional measurement actions but rather report the measurements already available to it. In the Multi-cell Round Trip Time (Multi-RTT) method, UE and the gNB both report time difference measurements to LMF to compute the position estimate. In Uplink Time Difference of Arrival (UL-TDoA) and Uplink Angle-of-Arrival (UL-AOA) methods, multiple gNBs measure the transmission time difference from the UE and angle-of-arrival based on the beam the UE is located in respectively, and send the measurements to LMF. Similarly, in Downlink Time Difference of Arrival (DL-TDOA) and Downlink Angle-of-Departure (DL-AoD) methods, the measurements are done on the UE side. The measurement results are sent to LMF for further computation. In Network-assisted GNSS (A-GNSS), WLAN, Bluetooth, Terrestrial Beacon System (TBS), and sensor-based positioning methods, the UE takes measurements of different wireless signals or sensors and sends the results to LMF for network side positioning, or alternatively, the UE can make use of the measurements, and optionally assistance data from LMF, to calculate the position itself.

3GPP in its specification [9] has envisioned the positioning accuracy needs for different vertical industries. For example, for the commercial handheld UE use case, the horizontal accuracy is 1-10 meters while vertical accuracy is under 3 meters with over 80% availability. Beyond these conventional positioning solutions, Machine Learning (ML) aided positioning techniques have been proposed in the literature (see [7] for a survey) to further improve the accuracy to the meter or sub-meter level.

Mobile Tracking Relevance. We make the following observations on 5G positioning:

- 5G cellular networks have the capability to localize a UE with a 10 meters accuracy. This is an address or sub-address level positioning precision, which reveals a lot of information about the user. In the 5G mobile phone use case, this localization accuracy is enough to enable mobile tracking by MNOs.
- MNOs can conduct UE positioning using measurements only from 5G-RAN, without the help of the UE (in 5G-RAN-assisted mode). It means that the user has no means

to stop non-intrusive positioning, thus loses control over the location privacy leakage to MNOs.

- 5G cellular networks operating in millimeter-wave frequencies will have ultra-dense access nodes deployed with inter-site distances ranging from a few meters (indoors) to 50 meters (outdoors) [26]. The small-cell 5G networks can localize UEs at the address level using cell IDs only. This enables mass surveillance of mobile users by MNOs, which further justifies the need for anti-mobile tracking solutions against MNOs.
- The energy efficiency introduced by Registration Area and the location privacy of the UE are two conflicting goals. The more power efficiency the UE can achieve, the more its movement patterns are leaked to the network from the RA configuration. We anticipate that the UE energy efficiency will degrade when anti-mobile tracking mechanisms are deployed. Disabling 5G RA altogether would enhance location privacy but also significantly increase user registration frequency. We recommend that a comprehensive trade-off evaluation should be conducted for future anti-mobile tracking solutions.

C. Various Levels of Mobile Tracking Threats

A cellular network *pinpoints* a user to collect a mobility event $< who, where, when, - >$, where who is a permanent identifier, $where$ is output from the positioning methods and $when$ is the time stamp. The MNO pinpoints the user continuously to get a set of mobility events. This data set can then be compiled to get the user's trajectory. The compilation of discrete events to trajectories is the process of user *tracing*.

The MNOs are also actively collecting user $activity$ data [27]–[29], which adds the last piece to the *mobility event* quadruple. $Activities$ include but are not limited to making phone calls, access to websites, and mobile app usage. Activities may disclose the social connections, online habits, and app engagement patterns of the user. Using the complete form of mobility event $< who, where, when, activity >$, MNOs have the ability to *profile* the users into structured data formats.

As we discussed in Section I, MNOs have the incentive to sell subscribers' location data. The good news is that legislative efforts are being made to ban the sale of customer location data [30], [31]. With these forthcoming regulations in mind, MNOs are trying hard to make a profit by selling *aggregated* mobile data. As an example, T-Mobile is grouping its Android subscribers into "personas" with user identifiers and location data concealed, and is selling the "personas" to advertisers by launching its new advertising platform [32]. Using these data, the marketers can extract intensive information about the users. Further, advanced tools like artificial intelligence (AI) can be used to do *inference*, *prediction* and even *interaction* [33] to improve the quality of LBS services. MNOs are more powerful than the marketers because they have the full data labeled with users' identifiers and locations. Thus, MNOs are able to conduct mobile tracking in a more general concept and pose more serious threats to subscribers' privacy.

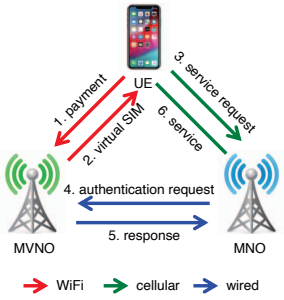

Fig. 4. ZipPhone [35] architecture.

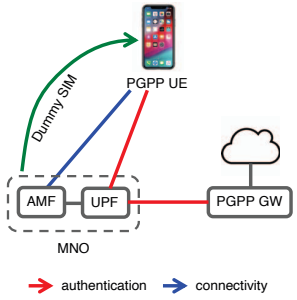

Fig. 5. PGPP [37] architecture.

IV. Anti-Mobile Tracking Solutions

Under the untrusted cellular network model, several solutions to counter the mobile tracking threats have been proposed in the literature, which can be categorized into identity anonymization and location obfuscation.

A. Identity Anonymization

Identity anonymization methods protect users' privacy by hiding users' identities from MNOs. In Section III-A, we have highlighted a list of permanent identifiers that can be used to link mobility events. Among them, SUPI is the main identifier used in 5G cellular networks (equivalent to IMSI in 4G). Most of the previous identity anonymization-based research works focus on nullifying SUPI/IMSI. While we focus on privacy protection against MNOs, readers are referred to [34] for a review on user identity privacy against non-MNO entities.

Ephemeral Identifiers. Sung et al. [35] introduced a new cellular service concept called ZipPhone that uses ephemeral IMSIs to break the association between the user and the IMSI by allowing users to buy virtual SIM cards over the Internet from Mobile Virtual Network Operators (MVNOs). As is shown in Fig. 4, ZipPhone does not have a physical SIM card. Instead, it connects to the Internet using side channels like WiFi to buy a virtual SIM card from MVNOs using anonymous payment methods. The user can then connect to MNO using the virtual SIM card, whose eligibility will be verified between the MNO and the MVNO.

In a newer version of ZipPhone [36], the authors expanded their previous work by identifying two attacks on user location privacy by cellular operators namely location profiling and trajectory linking. They modeled and quantified the attacker's accuracy when subscribers select from a set of identifiers and update the identifier frequently. However, ZipPhone has two limitations. Firstly, though the user has control of the update of his identifier, he has to sacrifice 5% uptime to improve privacy. Secondly, due to the user-selected identifier, a phone call over cellular networks is disabled because there is no call routing mechanism. Instead, the authors resort to Voice over IP (VoIP), whose service quality is not competitive with Voice over New Radio (VoNR) which uses dedicated 5G core networks.

Dummy SUPI. Schmitt et al. [37] proposed a hybrid phone privacy-preserving framework called PGPP using an identical

dummy SUPI for all subscribers to achieve unlinkability in cellular access. The system architecture is shown in Fig. 5. PGPP decouples the network connectivity function from authentication and billing. Each subscriber in the network is provisioned a SIM card with an identical SUPI, which is used to set up basic connectivity to the cellular network. A proprietary PGPP gateway is deployed on the user plane which is responsible for authentication and billing. The gateway assigns dynamic IP addresses to subscribers for data connections. This method has the same limitations as ZipPhone—the native phone call is disabled. Further, the solution is incomplete in terms of privacy protection because methods to eliminate linkability by other permanent identifiers (like GPSI and PEI) are missing.

B. Location Obfuscation

Ephemeral Tracking Area List. PGPP [37] also introduces a location obfuscation method targeting the paging procedure. It employs the concept of Tracking Area List (TAL) introduced in 4G LTE networks, which is defined as a continuum of a user's recent Tracking Areas (TAs). TALs are normally precomputed and assigned to UEs in existing networks. But in PGPP, TAL is generated on-the-fly by selecting a random number of adjacent TAs to provide improved location obfuscation among a larger group of mobile users. The idea of PGPP TAL is straightforward, but it increases signaling overhead as the network has to broadcast to a larger set of base stations.

P2P Cloaking. Tomasin et al. [38] proposed an integrated solution that includes three components: a virtual private mobility network (VPMN), a privacy-aware ecosystem, and a user awareness and control component. The latter two are introduced with a high-level perspective in the paper without sufficient details, so we classify this work as location obfuscation based on the main component VPMN. VPMN is a set of UEs working on device-to-device (D2D) mode, one of which is the relay of other UEs' communication between the cellular network and the VPMN. Following the previous work, Tomasin et al. [39] further considered a VPMN with multiple gateways. Naive selection of the gateways by UEs may create an unbalance of data rates at the gateways, which risks the disclosure of UE locations. This work adds a constraint of equal data transfer from each UE to all gateways as a remedy.

This idea of location cloaking using P2P connections is not new. Solutions following this thread have several limitations. First, the use of D2D communications mean that the UEs have to be mutually trusted, which is hard to achieve in the current cellular landscape. Second, the quality of the multi-hop D2D communication channels is not guaranteed due to UE mobility, which will inevitably introduce extra delays and packet loss, and degrade the end-to-end network performance. Third, reducing the location estimation accuracy from several meters to tens of meters does not essentially mitigate the mobile tracking threat posed by MNOs, who can conduct user profiling at address-level positioning accuracy.

V. Challenges and New Directions

Anti-mobile tracking as a means to preserve user identity and location privacy does not come without a cost. In this section, we raise several outstanding issues related to anti-mobile tracking that were overlooked in the literature.

A. Fine-grained Privacy Control by Users

An anti-mobile tracking scheme boils down to protecting the identity privacy and location privacy of mobile users. There have been numerous definitions of privacy because privacy is a subjective concept and different entities have diverse opinions on what privacy means to them. We believe that at the core of privacy conceptualization are *user awareness* and *user control*. Users should be able to know and determine when, how and to what extent information about them is shared with whom [40]. So, the first raised challenge is how user awareness is achieved. A practical privacy-preserving mechanism should include models to quantify user privacy leakage and to present the quantification results in a user-friendly way. On the other hand, it should allow users to actively manage their privacy exposure to the cellular networks by transferring the control of users' data from operators to users. To achieve this, the permanent identifiers in cellular networks have to be nullified. Only randomized temporary identifiers should be used, and the users should be able to refresh the identifiers whenever necessary.

Hence we identify the first open problem—*how to enable fine-grained user awareness and user control on user identity and location privacy leakage to the operators?* Location privacy quantification frameworks have been proposed in the literature [41]. We believe these frameworks will inspire researchers to develop mechanisms that can assist privacy leakage evaluation and privacy control by users.

B. The User Accountability Challenge

In a practical cellular network, holding mobile users accountable for their own actions or relevant public incidents is of both commercial and societal importance. For example, a user who exceeded the maximum data allowance in a data plan can be subject to rate throttling by the MNO; a mobile user who engaged in criminal activities during cellular access (e.g., committing telecom fraud) can be subject to identity and location reporting when requested by the law enforcement; in some cases, the law enforcement may also need to continuously monitor the voice calls of a certain individual for public safety.

The above accountability issues are never a real challenge in the existing 5G networks, where the MNO can uniquely identify a subscriber by the permanent mobile identifier (e.g., SUPI) which easily associates with the real user identity. Most jurisdictions require proof-of-ID registration for SIMs, even for prepaid options [42]. As for the monitoring task of law enforcement, the 5G framework also specifies a Lawful Interception (LI) mechanism to allow a Law Enforcement Agent (LEA) to access private communications of mobile users via a dedicated interface to the 5G Core [43].

In an idealistic anti-tracking cellular system where a user's real identity is always hidden from the MNO, user accountability is seemingly contradictory to the very goal of user privacy. This dilemma is further complicated by the law enforcement's existing privilege in the system, i.e., LI. Hence we posit an open problem—*How to achieve anti-tracking cellular access and user accountability simultaneously while accommodating law enforcement operation?* We envision that the solutions would include new cryptographic protocols that provide users anonymous cellular access by default and realize accountability functions with the joint effort of the MNO and law enforcement.

C. The Incoming Call Challenge

When the anti-tracking feature is active, receiving a phone call becomes a particular challenge if the user hides his permanent identifier from the serving MNO (e.g., using anonymity schemes from Section IV-A or other temporary identifier schemes). When a caller tries to call a user (i.e., callee) by a permanent identifier, the caller's MNO would not know which temporary ID the callee is using or which cell the callee is currently in. In this case, the MNO will not be able to direct the call to the callee.

Hence we raise an open challenge—*how to enable incoming call service without the MNO knowing the callee's real permanent identifier and cell location?* Specifically, a trusted caller who only knows the callee's permanent identifier should be able to make the voice call when the callee is connected to an MNO with a random temporary ID. Voice quality is also an important challenge. Entirely out-of-band workarounds like VoIP would not be desirable by users since they yield worse call quality compared to the 5G-native voice service (e.g., VoNR). Instead, we stipulate the possibility of a hybrid solution that combines the native voice service with an out-of-band callee discovery mechanism. It remains a technical challenge to return the temporary identifier and serving network of the callee to the caller's MNO without any linkage to the callee's permanent identifier.

VI. Mobile Tracking by Other Constituent Systems

Mobile tracking threats by location-based services and OS vendors are briefly reviewed in this section. Interested readers can see the listed survey papers and references therein for further information.

A. Mobile Tracking by Applications

Numerous location-based apps, such as Google Maps, Yelp, and Apple Weather, have been developed to offer highly customized and personalized services. As we discussed previously, a mobile OS obtains user location information through the embedded GPS module [4]. Any location-based app can collect user metadata, including precise location and hardware identifiers, by simply requesting user permission and gaining access to the metadata via developer APIs. It is of particular privacy concern that a user's mobility data are aggregated on an arbitrary LBS server and used for profiling or further analysis without the user's consent or awareness [44]–[46].

B. Mobile Tracking by OS Vendors

Mainstream mobile OS vendors like Apple and Google also rank among the most notorious user data collectors [47]. Apple iOS and Android running with the Google ecosystem have the transparent view and high OS privilege over a mobile device through their dedicated OS components and built-in services. For example, a hidden file *"consolidated.db"* which stores users' location history, was found on iPhone 4 during the 3G era [48]. Offline finding (OF), which Apple introduced in iOS 13, 2019, has become the largest location tracking system based on crowdsourcing nowadays [49]. OF can detect offline devices using Bluetooth Low Energy (BLE) and report the location information to the device owners.

Google Apple Exposure Notification (GAEN) [50], developed by Google and Apple during the COVID-19 pandemic, has been working as an effective tool to conduct digital contact tracing in over 20 countries since 2020. The core functions of GAEN, such as the broadcasting of rolling proximity identifiers and the storage of exposure keys, are embedded in the OS layer due to the use of BLE technology.

Google also collects a device's information (phone number, IMEI, IMSI, MAC address, IP address, etc.) and shares them with the Google Play service [51]. These metadata are synchronized with the Google server approximately every 5 minutes, which allows fine-grained location tracking [52]. Other than that, they can be explored by the Apps with lifted privilege, either granted by innocent users or through kernel-level attacks. Without reverse engineering, a smartphone user will not be able to delete the OS kernel code of the corresponding functions. Turning off or uninstalling the related services cannot prevent the covert functions from running. Massive surveillance conducted by Apple/Google is potentially possible.

VII. Conclusion

This paper studies the mobile tracking threats posed by mobile network operators. Specifically, we define the attack model, identify the root causes, and analyze the cases where subscribers' privacy has been compromised. We review the anti-mobile tracking solutions proposed in the literature, highlight the key research challenges, and suggest new research directions. Protecting user privacy against MNOs is a very challenging problem. A practical anti-mobile tracking solution may demand a radical new design and must jointly consider user anonymity, network utility, user accountability, and user privacy awareness and control simultaneously.

Acknowledgement

This work was supported in part by the US National Science Foundation under grants CNS-1837519 and CNS-1916902, the Office of Naval Research under grant N00014-19-1-2621, the Army Research Office under grant W911NF-20-1-0141, the Virginia Commonwealth Cyber Initiative, and a gift from InterDigital.

REFERENCES

[1] U. Cisco, "Cisco annual internet report (2018–2023) white paper," *Cisco: San Jose, CA, USA*, 2020.

[2] *5G performance measurements*, 3GPP TS 28.552 V16.9.0, 2021.

[3] G. L. INC, "Cellular network-based location system," U.S. Patent US5 519 760A, 5 21, 1996. [Online]. Available: https://worldwide.espacenet.com/patent/search/family/023002431/publication/US5519760A?q=pn%3DUS5519760

[4] V. Primault, A. Boutet, S. B. Mokhtar, and L. Brunie, "The long road to computational location privacy: A survey," *IEEE Communications Surveys & Tutorials*, vol. 21, no. 3, pp. 2772–2793, 2018.

[5] R. Di Taranto, S. Muppirisetty, R. Raulefs, D. Slock, T. Svensson, and H. Wymeersch, "Location-aware communications for 5g networks: How location information can improve scalability, latency, and robustness of 5g," *IEEE Signal Processing Magazine*, vol. 31, no. 6, pp. 102–112, 2014.

[6] J. A. del Peral-Rosado, R. Raulefs, J. A. López-Salcedo, and G. Seco-Granados, "Survey of cellular mobile radio localization methods: From 1g to 5g," *IEEE Communications Surveys & Tutorials*, vol. 20, no. 2, pp. 1124–1148, 2017.

[7] F. Mogyorósi, P. Revisnyei, A. Pašić, Z. Papp, I. Törös, P. Varga, and A. Pašić, "Positioning in 5g and 6g networks—a survey," *Sensors*, vol. 22, no. 13, p. 4757, 2022.

[8] FCC, "Indoor Location Accuracy Timeline and Live Call Data Reporting Template — fcc.gov," https://www.fcc.gov/public-safety-and-homeland-security/policy-and-licensing-division/911-services/general/location-accuracy-indoor-benchmarks, accessed 15-Jul-2022.

[9] *Service requirements for the 5G system*, 3GPP TS 22.261 V16.14.0, 2021.

[10] J. Cox, "I gave a bounty hunter $300. then he located our phone," https://www.vice.com/en/article/nepxbz/i-gave-a-bounty-hunter-300-dollars-located-phone-microbilt-zumigo-tmobile, 2019, accessed 15-Jul-2022.

[11] "Scott et al v. at&t inc. et al. case no. 19-cv-04063-sk," 2019.

[12] "Fitness app strava lights up staff at military bases," https://www.bbc.com/news/technology-42853072, 2018, accessed 15-Jul-2022.

[13] GSMA, "Whitepaper: Access to mobile services and proof of identity," 2021, accessed 15-Jul-2022.

[14] C. Paget, "Practical cellphone spying," *Def Con*, vol. 18, 2010.

[15] Apple, "Location services & privacy," https://www.apple.com/legal/privacy/data/en/location-services/, 2022, accessed 15-Jul-2022.

[16] *Digital cellular telecommunications system (Phase 2+)*, 3GPP TS 23.003 V16.3.0, 2020.

[17] *Security architecture and procedures for 5G System*, 3GPP TS 33.501 V16.9.0, 2022.

[18] "Itu-t recommendation e.212," https://www.itu.int/rec/T-REC-E.212, accessed 15-Jul-2022.

[19] B. Hong, S. Bae, and Y. Kim, "Guti reallocation demystified: Cellular location tracking with changing temporary identifier." in *NDSS*, 2018.

[20] H. Khan and K. M. Martin, "A survey of subscription privacy on the 5g radio interface-the past, present and future," *Journal of Information Security and Applications*, vol. 53, p. 102537, 2020.

[21] *System architecture for the 5G System*, 3GPP TS 23.501 V16.6.0, 2020.

[22] "Itu-t recommendation e.164," https://www.itu.int/rec/T-REC-E.164/, accessed 15-Jul-2022.

[23] *5G System (5GS) Location Services (LCS)*, 3GPP TS 23.273 V16.4.0, 2020.

[24] *Protocol and procedures for Lawful Interception (LI)*, 3GPP TS 33.128 V16.5.0, 2021.

[25] *Stage 2 functional specification of User Equipment (UE) positioning in NG-RAN*, 3GPP TS 38.305 V17.0.0, 2022.

[26] R. Baldemair, T. Irnich, K. Balachandran, E. Dahlman, G. Mildh, Y. Selén, S. Parkvall, M. Meyer, and A. Osseiran, "Ultra-dense networks in millimeter-wave frequencies," *IEEE Communications Magazine*, vol. 53, no. 1, pp. 202–208, 2015.

[27] Verizon, "Verizon custom experience programs faqs," https://www.verizon.com/support/verizon-custom-experience-programs-faqs/, 2022, accessed 15-Jul-2022.

[28] T-Mobile, "T-mobile privacy notice," https://www.t-mobile.com/privacy-center/our-practices/privacy-policy, 2022, accessed 15-Jul-2022.

[29] AT&T, "At&t privacy notice," https://about.att.com/privacy/full_privacy_policy.html, 2022, accessed 15-Jul-2022.

[30] "Fcc proposes over $200m in fines for wireless location data violations," https://www.fcc.gov/document/fcc-proposes-over-200m-fines-wireless-location-data-violations, 2020, accessed 15-Jul-2022.

[31] "Sweeping legislation aims to ban the sale of location data," https://www.vice.com/en/article/4axydq/legislation-aims-to-ban-health-and-location-data-protection-act, 2022, accessed 15-Jul-2022.

[32] "T-mobile advertising solutions," https://www.t-mobile.com/advertising-solutions, accessed 15-Jul-2022.

[33] S. B. Wicker, "The loss of location privacy in the cellular age," *Communications of the ACM*, vol. 55, no. 8, pp. 60–68, 2012.

[34] M. M. Saeed, M. K. Hasan, A. J. Obaid, R. A. Saeed, R. A. Mokhtar, E. S. Ali, M. Akhtaruzzaman, S. Amanlou, and A. Z. Hossain, "A comprehensive review on the users' identity privacy for 5g networks," *IET Communications*, vol. 16, no. 5, pp. 384–399, 2022.

[35] K. Sung, B. N. Levine, and M. Liberatore, "Location privacy without carrier cooperation," in *IEEE Workshop on Mobile Security Technologies, MOST*. Citeseer, 2014, p. 148.

[36] K. Sung, B. Levine, and M. Zheleva, "Protecting location privacy from untrusted wireless service providers," in *Proceedings of the 13th ACM Conference on Security and Privacy in Wireless and Mobile Networks*, 2020, pp. 266–277.

[37] P. Schmitt and B. Raghavan, "Pretty good phone privacy," in *30th USENIX Security Symposium (USENIX Security 21)*, 2021, pp. 1737–1754.

[38] S. Tomasin, M. Centenaro, G. Seco-Granados, S. Roth, and A. Sezgin, "Location-privacy leakage and integrated solutions for 5g cellular networks and beyond," *Sensors*, vol. 21, no. 15, p. 5176, 2021.

[39] S. Tomasin and J. G. L. Hidalgo, "Virtual private mobile network with multiple gateways for b5g location privacy," in *2021 IEEE 94th Vehicular Technology Conference (VTC2021-Fall)*. IEEE, 2021, pp. 1–6.

[40] A. F. Westin, "Privacy and freedom," *Washington and Lee Law Review*, vol. 25, no. 1, p. 166, 1968.

[41] R. Shokri, P. Papadimitratos, G. Theodorakopoulos, and J.-P. Hubaux, "Collaborative location privacy," in *2011 IEEE Eighth International Conference on Mobile Ad-Hoc and Sensor Systems*. IEEE, 2011, pp. 500–509.

[42] GSMA, "Access to Mobile Services and Proof-of-Identity: Global policy trends, dependencies and risks," 2018.

[43] *Lawful Interception (LI) architecture and functions*, 3GPP TS 33.107 V17.0.0, 2022.

[44] R. Gupta and U. P. Rao, "An exploration to location based service and its privacy preserving techniques: a survey," *Wireless Personal Communications*, vol. 96, no. 2, pp. 1973–2007, 2017.

[45] H. Jiang, J. Li, P. Zhao, F. Zeng, Z. Xiao, and A. Iyengar, "Location privacy-preserving mechanisms in location-based services: A comprehensive survey," *ACM Computing Surveys (CSUR)*, vol. 54, no. 1, pp. 1–36, 2021.

[46] R. Gupta and U. P. Rao, "Investigating and devising privacy preserving approaches for location-based services," in *Intelligent Technologies: Concepts, Applications, and Future Directions*. Springer, 2022, pp. 129–148.

[47] "Lawmakers want ftc to investigate apple, google over mobile tracking," https://www.wsj.com/articles/lawmakers-want-ftc-to-investigate-apple-google-over-mobile-tracking-11656077945?mod=latest_headlines, 2022, accessed 15-Jul-2022.

[48] N. Bilton, "3g apple ios devices are storing users location data," *The New York Times, Published: April*, vol. 20, p. 2011, 2011.

[49] A. Heinrich, M. Stute, T. Kornhuber, and M. Hollick, "Who can find my devices? security and privacy of apple's crowd-sourced bluetooth location tracking system," *arXiv preprint arXiv:2103.02282*, 2021.

[50] Apple, "Privacy-Preserving Contact Tracing - Apple and Google — covid19.apple.com," https://covid19.apple.com/contacttracing, accessed 15-Jul-2022.

[51] "Learn about the Android Device Configuration Service - Android Help," https://support.google.com/android/answer/9021432, accessed 15-Jul-2022.

[52] D. J. Leith, "Mobile handset privacy: Measuring the data ios and android send to apple and google," in *International Conference on Security and Privacy in Communication Systems*. Springer, 2021, pp. 231–251.