ANOTEL: Cellular Networks with Location Privacy (Extended Version)

Tim Dittler Florian Tschorsch Stefan Dietzel Björn Scheuermann
Humboldt-Universität zu Berlin, Berlin, Germany
{dittler, tschorsch, scheuermann}@informatik.hu-berlin.de, stefan.dietzel@hu-berlin.de

***Abstract*—Location management is a key component of cellular networks. From a privacy perspective, however, it is also a major weakness: location management empowers the network operator to track users. In today's public and scientific discussion, the centralized storage of location data is mostly taken as a fact, and users are expected to trust the network operator. ANOTEL presents a novel, clean-slate approach of location management in cellular networks that challenges this assumption. We developed a design that is able to route calls to users who move through cellular networks, without violating their location privacy. We evaluate our approach using simulations and a practical user tracking algorithm.**

I. Introduction

Mobile cell phones have become ubiquitous location tracking devices. They register to nearby base station and expose their location in real time, facilitating spatio-temporal mobility profiles [1]. Everyone with access to location data—be it external attackers, government agencies, or even the network provider—can (mis-)use the data to track and predict individual user movements.

Surprisingly little literature questions this situation; it seems as if this privacy threat has been widely accepted. Previous work on location privacy mainly focuses on privacy towards location-based services (LBS) rather than on the architecture of cellular networks [2], [3], [4]. We, instead, identify and solve three major privacy issues of cellular network location management: a) location data and identifiers are stored in and controlled by a central entity, b) an attacker can force devices to update their location, which, in case of stealth pings, remains unnoticed by the user [5], and c) paging signals are plaintext broadcast messages and hence leak information about the callee's identity and location.

To this end, we design and evaluate ANOTEL, a novel, distributed approach to location management, which fundamentally rethinks the network provider's role. ANOTEL is a clean-slate approach, rooted in the privacy by design principle [6]. Accumulations of sensitive data are not merely protected by policies in our design, but they are avoided in the first place.

We intentionally do not tie our design to a specific cellular network standard. Rather, we focus on the mobility management core components that are required to instantiate any mobile network. Therefore, our generic approach can be applied to a number of specific network standards. We restrict the network operator to its core task of providing the cellular infrastructure. Further, we add a location provider and a pseudonym provider as independent entities and build efficient, privacy preserving call signaling on top.

This is a hard challenge, since no entity, including the network operator, shall ever provided with both the users' identity and location information. Users therefore remain anonymous, particularly towards the network provider. In situations where device localization is required to establish calls, it must not be possible to infer user identities from exchanged information. These aims require novel, non-obvious signaling procedures, in particular for paging and call setup, which we devise as one key contribution of this paper. We also show how billing and payment remain possible.

We present a simulation model that implements an attacker who uses a practical, distance-based user tracking approach. Our evaluation shows that such an attacker is unable to directly trace users in the network. ANOTEL significantly increases location privacy not only towards service providers, but also towards infrastructure operators. Therefore, ANOTEL protects from unnoticed collection of location data by third parties and therefore enables location data sovereignty for users.

Our main contributions can be summarized as follows: We present a distributed, anonymized location management scheme, which enables location and identity privacy towards the network operator. We integrate location management with enhanced paging and call signaling, so that an active attacker cannot generate valid localization signals. Moreover, we propose a practical way to evaluate location privacy in cellular networks and apply it in order to confirm our design.

Next, we explore related work in Sec. II. In Sec. III, we define our system and adversary model. Thereafter, we describe ANOTEL's architecture and specify internal and external call establishment in Sec. IV. We present our simulation and evaluation of ANOTEL in Sec. V, and discuss non-call features, such as billing and authorization, in Sec. VI. Sec. VII concludes this paper.

II. Related Work

There is a significant body of work on location privacy for location-based services [2], [3], [4], [7]. A primary focus is privacy towards a service provider operating within a generally trusted network infrastructure. In contrast, our intention is not to hide location profiles from a service provider but instead from the network infrastructure itself. Therefore, this research direction is orthogonal to our approach.

Work on WiFi location privacy has included network operators as potential attackers [8], [9], [10]. Focusing on a different class of networks, however, these works discuss neither location management nor call signaling.

Limiting the network operator's knowledge in cellular networks is often closely related to the specifics of GSM [11], [12], [13]. [11], [12] aim to host location data on user-operated devices or at a trusted third party. However, connecting to a specific device constitutes a fingerprint, which can be used to reduce the anonymity set and, ultimately, to identify users. While the use of mix networks, as in [12], mitigates the problem, it implies other weaknesses, for instance with respect to availability. [13] heavily depends on Tor [14] for pseudonym changes and for routing incoming calls. ANOTEL prevents from fingerprinting and tracking: location data is fully decoupled from identities. Moreover, no additional user-operated components or external anonymity services are needed.

[15] suggests to distribute user pseudonyms between n parties. Location data is still stored centrally, but can only be linked to pseudonyms. An attacker may, however, still generate incoming traffic and, thereby, trigger a pseudonym resolution. In ANOTEL, pseudonyms are implemented as random one-time handles, and network operations, including call routing, do not need to resolve pseudonyms.

ANOTEL is designed with a global attacker in mind who has the same knowledge as the network operator. Although encryption of traffic is a necessary component, its use alone, as proposed in [16], does not suffice to thwart such an attacker. Unlike in [17], it is also not possible to gain privacy by hiding information from base stations.

Research on paging protocols focuses mostly on efficiency [18], [19], [20]. [21] presents a paging strategy that incrementally improves location accuracy using sequential pagings. The approach avoids location storage but still reveals the callee's position during every call setup. [22] proposed a modified paging scheme that encrypts user identifiers in physical tags, which are only recognizable by the user herself. While this approach offers good protection against malicious users who monitor paging messages, it may degrade the user experience when the signal-to-noise ratio is too low. ANOTEL achieves similar privacy gains by using one-time pseudonyms without the need for physical layer modifications.

[23] criticizes the infrequent change of temporary mobile subscriber identities (TMSIs) in today's cellular networks, which facilitates tracing users. Location privacy, though, depends on not knowing the association between long-term identity (IMSI) and TMSI. In traditional networks, the TMSI-IMSI association is always known to the network operator. Additionally, anyone with access to SS7 can directly resolve TMSIs and attack users' location privacy [24]. ANOTEL precludes such attacks by avoiding knowledge of user locations in the first place. We introduce one-time paging names, which improve on TMSIs. No single entity can resolve ANOTEL's paging names to identities, and our paging protocol does not require such kind of resolutions.

With increasing attention on privacy, clean-slate approaches became more and more popular, mostly focusing on future Internet designs [25]. ANOTEL, too, is a clean-slate design, but focuses on cellular networks. We achieve location privacy against a global attacker who monitors all network access.

III. SYSTEM AND ATTACKER MODEL

For the discussion of our protocol design, we consider an abstraction of cellular phone networks. We assume a network that is (much) larger than a single cell's coverage area. Therefore, location information is needed to route calls to the called user's current cell. The network is operated by a party known as *network operator* and clustered into *location areas (LAs).* Each LA consists of one or more *cells,* each covered by one *base station.*

In today's cellphone standards, the users' current locations are stored in *location registers* controlled by the network operator. These are known under varying names, such as home/visitor location register in GSM/UMTS/CDMA2000 and mobility management entity in LTE. In all cases, they map globally unique user identifiers to locations and are accessible to the network operator.

When a user enters a new LA, she sends a *location update* to inform the network of her new location. If a user is to receive a call, a paging message is broadcasted within the corresponding LA. When a user establishes a connection to a base station, the accuracy of her location is updated from LA level to cell level.

In this paper, we raise the fundamental question: *How to protect location privacy against an adversary who shares the view of the network operator?* We assume that a semi-honest attacker [26] has global network knowledge. That is, the attacker follows protocols but stores all accessible data and tries to infer additional knowledge about the users. We consider attackers successful if they are able to build spatio-temporal user movement profiles. In traditional cellular networks, the above described attackers can trivially build profiles using message identifiers called TMSI. Such direct tracking serves as our baseline.

We do not consider attacks that require collaboration between ANOTEL's major entities. Rather, we assume that an attacker is unable to gain control over multiple distinct entities. Collaboration attacks would trivially subvert the separation of concerns within ANOTEL. As a result, location privacy would be reduced to the level currently implemented in cellular networks.

IV. ANOTEL

We propose ANOTEL as an alternative to traditional cellular network location management architectures. In this section, an outline of ANOTEL's system design and key entities will be followed by an in-depth discussion of its call signaling.

The network operator N remains a central entity in our design. N manages the mobile base stations and operates the backbone network. Therefore, N can observe all traffic, including its origin and destination. User identities, however, are never revealed to N.

In ANOTEL, incoming calls from external networks are not directly handled by N. Instead, they are routed through a separate entity, the *pseudonym provider P*. P maintains a mapping between a long-term user ID (including phone numbers) and a temporary pseudonym for each user. Only these pseudonyms are used in the core network and are visible to N. While P knows the mapping between pseudonyms and user IDs, P never gains any information about user locations.

The current locations of users are stored at a third entity, the *location provider* L, who maintains a mapping from temporary pseudonyms to locations. L is, however, not able to translate the pseudonyms to user identities.

The separation between L and N is a core feature of ANOTEL's design. Had we only introduced P, N could have resolved pseudonyms by initiating calls and observing paging signals issued by P. In ANOTEL, N is not able to observe the currently paged pseudonyms. To maintain privacy, it is crucial to carefully design user paging: on one hand, efficient call routing is required to signal incoming and outgoing calls, on the other hand, such routing reveals user locations in traditional designs. Therefore, ANOTEL's novel paging protocol is a key ingredient to achieve user privacy.

In the following, we describe each component's role in more detail, focusing on ANOTEL's signaling protocol as a key enabling factor of user privacy.

A. Keys and Pseudonym Management

We assume that the pseudonym provider P shares cryptographic keys with the network operator N and with the location provider L. We assume the existence of shared symmetric keys $K_{P,N}$ and $K_{P,L}$, which could, of course, be negotiated using any standard key exchange protocol. Furthermore, P and L each have asymmetric encryption key pairs; the matching public keys K_P and K_L are assumed to be known by all users. In the following, $\{x\}_{K_{y,z}}$ denotes that a message x is encrypted using a symmetric key $K_{y,z}$ shared between y and z. $\{x\}_{K_y}$ denotes asymmetric encryption of message x using y's public key. The employed encryption scheme should be probabilistic and non-malleable.

Each user has a long-term identity id, comparable to an IMSI or phone number. The user's data is stored under a temporary pseudonym p. Pseudonym changes are performed frequently: upon every location update, each outgoing call, and also time-based after a too-long period without a pseudonym change. Assume Alice (A) is an ANOTEL user. To update her pseudonym, A chooses a new pseudonym p at random. Then, A transmits p to P in a *pseudonym update* message $A \rightarrow P\colon \{id, p, k, t\}_{K_P}$. The timestamp t ensures freshness; k authenticates A towards P and may be transmitted using a more sophisticated password scheme (e. g., [27]) in practice.

Like any cellular network, ANOTEL needs location update messages in order to keep track of the users' locations. This information is outsourced to L, though: users send location updates to L when they change LAs. Alice knows her current location area l from periodic announcement beacons sent by the base stations. The update message contains A's pseudonym p, the current location area l, and a random seed s, which will be used for paging as discussed below. The pseudonym p and seed s must not be readable by N. l, in contrast, is anyway known by N, who operates the base station. Messages, where l and the receiving base stations do not match, will be ignored. The complete message format is $A \rightarrow L\colon l, \{p, s\}_{K_L}$, where p is A's current pseudonym. Note that L does not know to which user the pseudonym belongs. For this reason, L is not (and should not be) able to correlate subsequent pseudonyms of the same user. Further, L cannot observe how long a pseudonym remains in a specific location. This design also implies that L cannot authenticate pseudonym updates. Such authentication is actually not necessary in ANOTEL: the pseudonym is random and cannot be guessed by any instance but L or P. Pseudonym changes trigger database updates at L, mapping the new pseudonym to the current location, without explicit removal or substitution of the previous pseudonym's mapping. An entry created by a location update is protected for a certain time span and automatically deleted after. If a user stays in a location for longer than the lifetime of her pseudonym, she needs to refresh her records by triggering a time-based update. To prevent pseudonym correlation, time-based updates also create new, unrelated records rather than replacing existing ones. To harden against attacks that exploit known update rhythms, time-based updates are performed at a random point in time within the second half of the location record's lifetime.

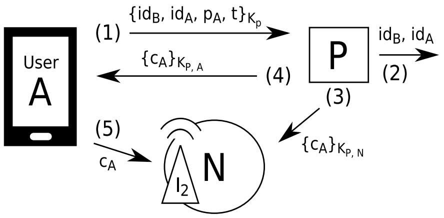

Fig. 1. Outgoing call signaling (i. e., from ANOTEL to a foreign network).

B. Call Establishment

Call setup, including the necessary routing and paging steps, is a key factor for a cellular network's privacy properties. Next, we describe our design's main elements using a detailed description of outgoing and incoming calls. We also discuss the amount of knowledge that each entity gains, and thereby show how user privacy is realized. To simplify the explanations, we focus on the signaling of a single call between two users.

1) Outgoing calls: Alice (A) is an ANOTEL user, currently located in location area l_2. She wants to call Bob (B), who uses a foreign network. First, Alice needs to prove that she is an authorized network user. Instead of the traditional approach of using her identity to do so, we introduce anonymous tokens generated by P. Tokens act as a one-time secret between P, N, and A. The token confirms towards N that Alice is a valid call endpoint and that her identity has been checked by P.

Fig. 1 shows our protocol step by step. First, Alice signals to P her intention to call Bob:

$$A \rightarrow P\colon \{id_B, id_A, p_A, t\}_{K_P}. \qquad \text{(step 1)}$$

The message includes her pseudonym p_A and a timestamp t. P is located on the cellular network's edge and as a signaling proxy for the users, which can issue and receive legacy calls without N being able to see the involved ids. A and P establish a connection via N's network and keep it open until the signaling is finished. N uses this connection to route the messages between P and A without knowing A's identity.

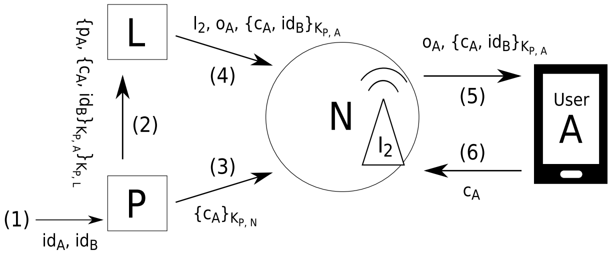
L
P
N
User
A
l_2
(1)
(2)
(3)
(4)
(5)
(6)
id_A, id_B
$\{p_A, \{c_A, id_B\}_{K_{P,A}}\}_{K_{P,L}}$
$\{c_A\}_{K_{P,N}}$
l_2, o_A, $\{c_A, id_B\}_{K_{P,A}}$
o_A, $\{c_A, id_B\}_{K_{P,A}}$
c_A

Fig. 2. Incoming call signaling (i. e., from a foreign network to ANOTEL).

After receiving Alice's message, P checks t for freshness (within a few seconds) and validates Alice's id and pseudonym. If checks succeed, P signals Alice's call to the foreign network F:

$$P \to F\colon id_B, id_A. \tag{step 2}$$

Inter-network signaling is the same as in regular networks. If signaling succeeds, P generates an anonymous token c_A from a sufficiently large field, so that it does not collide with other tokens and cannot easily be guessed. P then sends the (encrypted) token to N and also to Alice, via the connection established in step 1:

$$P \to N\colon \{c_A\}_{K_{P,N}}, \tag{step 3}$$

$$P \to A\colon \{c_A\}_{K_{P,A}}. \tag{step 4}$$

Finally, Alice reveals c_A to N:

$$A \to N\colon c_A. \tag{step 5}$$

The token c_A proves to N that Alice is a valid caller without revealing her identity. At the same time, the token indicates to which call at P Alice should be connected via N's network.

In contrast to traditional implementations, N neither learned that Alice and Bob are talking nor where Alice is located—only that "some user" in l_2 has established a call. P knows Alice's identity but not her current location. Both N and P are thus unable to create individual movement profiles.

The timestamp in Alice's first message prevents replay attacks by an adversary. This feature requires loosely coupled clocks in the network, but in turn avoids additional per-user state (such as sequence numbers) and additional round trips (such as for a challenge-response mechanism). Preferably, the network only allows one call per timeframe-and-user tuple. Moreover, an attacker can never establish a call under Alice's identity without knowledge of the token c_A. Once a token has been revealed in step 5, it cannot be reused.

2) Incoming calls: Incoming calls are more complex than outgoing calls: they require location management. Assume that foreign network user Bob (B) wants to call ANOTEL user Alice (A). A resides in location area l_2, as depicted in Fig. 2.

First, the foreign network signals Bob's call request to P:

$$F \to P\colon id_A, id_B. \tag{step 1}$$

P translates id_A to A's current pseudonym p_A and waits for the next paging cycle, in the meantime collecting further incoming call requests.

The waiting time is a key element in ANOTEL: it mixes incoming calls and prevents L from learning pseudonym-user mappings. The paging frequency should be chosen so that it mixes several calls without impacting the user experience; thus, mixing is easier for networks with high call frequencies. We hypothesize that waiting one second provides sufficient privacy in many cases.

In the next paging cycle, P sends a message to L to initiate paging. The message structure for Bob's call request is

$$P \to L\colon \{p_A, \{c_A, id_B\}_{K_{P,A}}\}_{K_{P,L}}. \tag{step 2}$$

c_A is a randomly chosen anonymous call token, as discussed for outgoing calls. Analogous messages are created for all other incoming calls for the same paging cycle; they are transmitted in random order within a single list of messages.

The pseudonym provider also informs N about Bob's call:

$$P \to N\colon \{c_A\}_{K_{P,N}}. \tag{step 3}$$

Recall that during the location update, Alice transmitted a seed s to L. This seed s is used to compute one-time paging names o_i ($i \in \mathbb{N}$) as $o_1 = h(s)$ and $o_{i+1} = h(o_i + s)$, where h is a cryptographic hash function, such as SHA-256. Whenever L needs to page A, L computes a fresh o_i. Only L and Alice can compute her current paging names. In particular, A's paging names are unlinkable for an attacker. At the same time, the hash-based generation of o_i is computationally efficient for both L and A.

One-time paging names require synchronization between L and A. Since there is no transmission reliability for paging, A may miss pagings due to bad reception. She would then listen for an outdated paging name and would miss subsequent pagings. To solve this problem, A does not only listen for her presumed current paging name, but for a list of consecutive paging names. If she recognizes one of them in a paging, she discards all previous paging names and pre-computes new ones based on the most-recently received paging name.

Having computed Alice's current paging name o_A, L constructs the paging signal for the corresponding location l_2: $l_2, o_A, \{c_A, id_B\}_{K_{P,A}}$. This process is repeated for every incoming call in P's list. For calls in the same location, the paging information is appended to the existing packet. The result is a list of pagings with one packet per LA. These paging packets are subsequently filled with dummies, i. e., with paging messages that address non-associated, random paging names. Dummies prevent observers from building usage profiles based on existence and size of paging packets for different LAs.

The location provider L hands over the paging packets—one per LA—to N. For Alice's location, the packet is:

$$\begin{aligned} L \to N\colon l_2,\ & o_A, \{c_A, id_B\}_{K_{P,A}}, \ldots, \\ & o_m, \{c_m, id_{caller_m}\}_{K_{P,U_m}}. \end{aligned} \tag{step 4}$$

Here, m denotes the maximum number of pagings per LA and paging cycle. For example, $m = 56$ pagings in a 1280 ms period for LTE networks [28]. The overhead incurred by padding is negligible in relation to the capacity of a cell in today's cellular networks. Moreover, the overhead shrinks as the number of active users increases.

From N's perspective, all packets have the same size, and their content is unknown to N. Likewise, an eavesdropper in the wireless network cannot distinguish paging packet contents. Without knowledge of Alice's currently used paging name, the paging messages do not reveal her current location.

The network operator N broadcasts the paging packets within their designated location areas:

$$N \rightarrow U_1, \ldots, U_n:\ o_A, \{c_A, id_B\}_{K_{P,A}}, \ldots, o_m, \{c_m, id_{caller_m}\}_{K_{P,U_m}}. \qquad \text{(step 5)}$$

Here, n is the number of all users U in the location area. Alice's phone listens for paging broadcasts in her current location area and checks if any one-time paging name o_i is among her currently used set. If a paging name is recognized, the phone decrypts $\{c_A, id_B\}_{K_{P,A}}$. Now the phone rings to inform Alice about the call, and it shows the identifier of the caller Bob. In contrast to traditional cellular network designs, no session between N and A has yet been established.

Only if Alice decides to answer the call, she publishes c_A, authenticating herself as the call endpoint:

$$A \rightarrow N\colon c_A. \qquad \text{(step 6)}$$

After Alice answered the call, N knows that someone from a foreign network is calling a user in a cell in l_2, but N knows neither the identity of the caller nor that of the callee. P knows that Alice and Bob talk, but not where they are. L knows that someone in l_2 was paged, but not who and whether the user answered the call.

Internal calls between ANOTEL users, are implemented using a straightforward combination of incoming and outgoing call signaling as described above. That is, in essence, they are established via P.

In summary, the major distinctive feature of ANOTEL is the separation between network operator N, location provider L, and pseudonym provider P. We have shown how our system design uses these entities to decouple user identities from location management and network operation, strengthening user privacy towards N. We will now proceed with an evaluation of the location privacy achieved with the proposed mechanisms, before we subsequently turn to a discussion of deployment aspects including, among others, the handling of mobile data connections and billing.

V. Evaluation

ANOTEL is designed to protect users' location privacy against a global attacker, which has the same knowledge and capabilities as the network operator. Notably, the attacker can observe all network communication.

We assume that an attacker, who has acquired a single position of a user, wants to trace her movement. The attacker tries to link consecutive connections and thereby construct a trace. Such tracking is trivially possible in today's cellular networks, because up-to-date location information is regularly sent to the network operator. In ANOTEL, tracking attempts are more challenging. Therefore, we measure ANOTEL's *gain* in user privacy using the *reduction* of tracking success, which is a common privacy quantification approach [29], [30], [31].

We model the cellular network in a simplified way using a grid of cells, which are randomly grouped into rectangular LAs. Our scenario is supposed to resemble a typical rush hour in Berlin, Germany. We use the base station data of OpenCellID [32] from 2014 to parametrize the grid construction. In particular, we extract the number of cells and LAs for UMTS (for which the data is both extensive and recent) and use these as a basis. On average, each of the four major network operators manages around 2500 UMTS cells and 42 LAs in Berlin [33].

TABLE I
Capabilities and Knowledge of Anotel's entities.

		U	P	L	N
Capabilities	$id \leftrightarrow p$	●	●		
	$o \rightarrow id$	●			
	$p \rightarrow l$	●		●	
	$p \rightarrow o$	●		●	
	issue calls	●	●		●
Knowledge	participants	●	●		
	source cell	if caller			●
	destination cell	if callee		●	if accepted
	time	●	●	●	●
	duration	●			●

Based on this network model, we implemented a discrete event simulation, which simulates user movement and call events. During the simulation, users move between randomly chosen pairs of origins and destinations at constant speed. We choose users' individual velocities such that their journeys take the whole simulation time, i. e., 120 minutes. Calls are randomly and uniformly distributed over this time. During simulation setup, we randomly choose a number of users and make their initial location known to the attacker.

As an example, Fig. 3 illustrates a 20-user scenario where each user places six calls, which corresponds to intensive cellular network usage. The results clearly show the fundamental privacy challenge: without protecting measures, attackers can trace users by simply listening to infrastructure messages, i. e., location updates, paging, and call establishment.

In ANOTEL, an attacker cannot simply eavesdrop on location updates. Table I summarizes the capabilities and the knowledge of ANOTEL's entities. Attackers with the knowledge of N cannot decide to which user identity a message belongs, because they cannot translate one-time paging names o to pseudonyms p or to ids. Attackers are therefore incapable of linking consecutive connections to a single user. Since attackers cannot distinguish genuine pagings from dummies, unanswered incoming calls leak no location data. The same is true if we assume that L is compromised: it does not know which user a call belongs to. Since users employ different pseudonyms for each LA, L cannot link consecutive positions. While attackers with the knowledge of P could link pseudonyms, they gain no information about any locations and can, therefore, not trace any movements.

From N's and L's perspectives, there is no direct link between consecutive pseudonyms of the same user. As a result, any attempt to heuristically track user movements results in ambiguous mappings. Assume that a new pseudonym shows up, and an attacker attempts to identify its predecessor to trace user movement. The attacker must at least take into account as predecessor candidates all pseudonyms that have occurred within the same or a neighboring LA in the recent past. Moreover, a user may have switched her device on or off, introducing further tracking mismatches. This effect is comparable to mix zone approaches [2], but instead of implementing dedicated mix zones where users have to remain silent, ANOTEL uses the silent times produced by passive connectivity to continuously remix identities. In a sense, ANOTEL is designed in such a way that *every* cell has a mixing effect and, hence, prevents linkability.

Therefore, even attackers with very good heuristics will be able to match two consecutive pseudonyms of the same user with a probability of at most $1-\epsilon$, for some $\epsilon > 0$. Not having lost track of the user after n pseudonym changes is therefore only possible with probability $(1-\epsilon)^n$. That is, the probability of successful tracking over longer time spans converges to zero. Convergence will be faster when user density is higher, as matching errors are more likely in high density scenarios.

To quantify ANOTEL's privacy gains in a specific scenario, we implement a matching heuristic and apply it to our Berlin scenario. Namely, whenever a new signal occurs, the attacker appends it to the user for whom the previously assumed position is closest to the location in the new signal. Our heuristic, together with the scenario and evaluation methods, is publicly available [33]. Our distance-based matching heuristic leverages movement patterns commonly found in cellular network usage, which is a benefit over multi-hypothesis tracking (MHT), a method often applied in related work [29], [31]. In MHT-based approaches, users are assumed to keep speed and direction with higher probability than to change them. However, MHT fails to detect back-and-forth trips, which are common in cellular networks. Also, closely spaced targets are usually a problem for MHT solutions. Since the spatial resolution in our network is low compared to other tracking methods such as GPS, this limitation applies heavily.

To investigate ANOTEL's mixing effect in detail, we vary the number of users and the number of events. Fig. 4 illustrates the impact of the number of users on the fraction of correctly matched point sequences. We call the latter the attacker's success rate. This success rate is calculated for sub-paths of length two and three and for the complete paths, which have varying length. Error bars show the 95% confidence interval of 50 simulation runs (most errors are very small and thus hardly visible in the figure).

The higher the number of current users in a cell, the less likely the attacker is to guess correctly which user moved to another cell. If there is only one user (and if this fact is known to the attacker), the attacker can—obviously—track her with complete certainty. All pseudonyms must belong to this user. With an increasing number of users, the success rate drops rapidly. Also, it is harder for the attacker to trace longer paths, because of the above-mentioned mixing effect. The remaining tracking success is mainly due to users who are located in sparsely populated areas and are, therefore, easier to follow. As Fig. 4 shows, the attacker's tracking heuristic will completely fail when there are more users than cells. Because users are identified by their start cells, there is no way for an attacker to distinguish two users in the same cell.

Tracking success is also influenced by the number of events. It is easier to trace users if they communicate more frequently. Fig. 5, however, shows that this effect becomes negligible with a growing number of users. That is, the impact of the number of users (x-axis) on the success rate (y-axis) is much stronger than the influence of the events per user (z-axis).

In edge cases, where an attacker knows with (unrealistic) certainty that only one user occupies a larger area, tracking opportunities in ANOTEL are comparable to those in today's cellular networks. Therefore, we further investigate the expected number of users, cells, and LAs in Berlin. In its inner city, Berlin is inhabited by about 1.7 million people. As 90% of Germans own a cellphone [34], there are, on average, 612 cellphones per cell. According to [35], an average American adult sends or receives 42 messages and 12 calls per day. Numbers for Germany, which are not available, are likely to be comparable. Since every event is counted at incoming and outgoing participant, these numbers amount to $\frac{42+12}{2} = 27$ paging messages per day and person. Thus, an attacker would likely see more than ten pagings per second and area:

$$612\ \frac{\text{phones}}{\text{cell}} \cdot \frac{2500\ \text{cells}}{42\ \text{areas}} \cdot 27\ \frac{\text{pagings}}{\text{phone} \cdot \text{day}} \approx 11\ \frac{\text{pagings}}{\text{area} \cdot \text{sec}}.$$

As discussed before, the attacker heuristic described above is likely to fail under these conditions.

In summary, we have shown that, in traditional networks, an attacker with access to the location register can harvest all users' location data in real-time. In ANOTEL, the mixing of pseudonyms hinders an attacker from launching such attacks. An adversary would have to utilize an advanced heuristic with extensive logic, additional knowledge, more computing power, and larger storage space to trace even some users. This contrast shows that ANOTEL significantly increases location privacy compared to existing cellular networks.

VI. Deployment Challenges

In this paper, we focus on core components of cellular networks and present a design for their distributed, privacy-preserving implementation: location management and call establishment. To underline ANOTEL's concepts, we refrain from intertwining its design with existing cellular network specifications. We emphasize, however, ANOTEL's practical applicability and discuss in the following, how its design can be extended to address common deployment challenges.

A. Text Messages

Text messages are critical to location privacy, since incoming messages usually produce location samples. In ANOTEL, sending text message is analogue to signaling an incoming call. The message is sent asymmetrically encrypted to the pseudonym provider P. The pseudonym provider decrypts the message and builds a paging signal for the recipient. But instead of a token c, P encrypts the message in one or more paging signals. Since the transmission medium is unreliable, the user should acknowledge messages the next time she communicates with P. The delay of acknowledgments avoids unnecessary connection establishments.

In contrast to traditional cellular networks, ANOTEL prevents eavesdropping on text messages by applying cryptography to every message by default. In particular, N and other

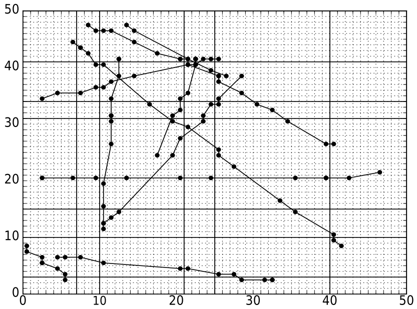

Fig. 3. Tracking in a traditional cellular network (dashed/strong lines represent cell/LA borders; points denote call events).

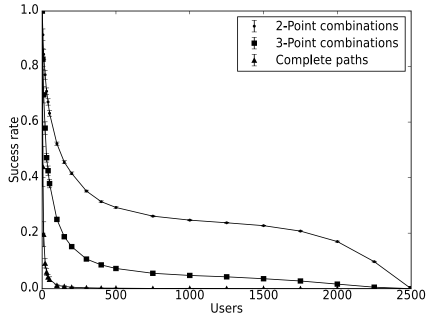

Fig. 4. Tracking results for six events with varying event times, user paths and location areas (50 runs, 95 % confidence intervals).

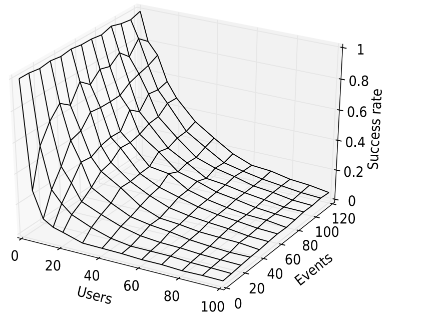

Fig. 5. 3D wireframe plot showing the impact of the number of users and events on the success rate (complete paths, mean of 50 runs).

users cannot read messages. Of course, since P processes the plaintext, additional end-to-end encryption is still advisable for message integrity and privacy.

B. Data Connections

ANOTEL is able to grant anonymous Internet access. In contrast to anonymity networks like Tor [14], however, ANOTEL does not build an overlay over existing infrastructure. Instead, it proposes a new infrastructure. Thus, ANOTEL can be considered a privacy-enhancing link layer below the TCP/IP protocol stack.

With every pseudonym update, the user receives a set of IPv6 addresses from N. Since N cannot link pseudonyms, N has no means to assign static IP addresses. It is important, though, that the IPs should not be of longer lifetime than the pseudonyms. Therefore, users should regularly change IPs. Ideally, different online activities should use different IPs to make correlations more difficult, as also suggested in IPv6's privacy extensions [36]. In addition, it is conceivable to use multiple pseudonyms in the same LA to distribute activities also over different pseudonyms.

C. Billing

Mobile phone plans with periodic membership fees or flatrates are commonly used in traditional networks. ANOTEL trivially supports these solutions: to facilitate such a plan, we introduce a bank as an additional entity for payment handling. The bank activates the user's id according to her plan. All services can then be used as described before, with revenues shared between N, L, P and the bank.

In addition, ANOTEL's design also supports a prepaid-like, privacy-preserving payment scheme. A centralized anonymized electronic cash system, similar to [37], can be used for accounting. As with regular pay-as-you-go approaches, users top up in advance, but receive virtual "coins", which represent their balance. The coins are issued by the bank and carry a blind signature [38], i. e., a digital signature of the bank without any association between the coin and the user. Thus, neither the bank nor any other entity can re-identify a user based on the appearance of a specific coin. Blind signatures yield a verifiable, forgery-proof, and anonymous electronic cash system. In order to pay for a call, the callee may attach coins to the payload, which the network operator can verify and redeem at the bank while the identity of the user remains unknown. Post-paid plans can be emulated based on the mechanisms described above.

Using the described approaches, the bank constitutes another entity with separated concerns alongside N, P and L. In addition, using an anonymous payment scheme renders coalitions with the bank worthless in terms of attacks on privacy. This underlines that ANOTEL's strong privacy properties can be achieved within a commercial environment. It also provides a basis for rewarding N, P, and L for their respective services, without giving up anonymity.

D. Authentication

Most deployed authentication schemes for cellular networks, such as EPS AKA for LTE [39], are not suitable ANOTEL's design goals. EPS AKA, for instance, provides mutual authentication between a user entity and a respective home subscriber server (HSS). Since user credentials are centrally stored in the HSS, authentication implies linking with a long-term identity, which would enable tracking of individual users. While the use of electronic cash for authentication to preserve privacy is possible [12], a DoS attack targeting the network operator, for example, would force users to re-authenticate and thus lose money.

We therefore use *periodic n-times anonymous tokens* [40] for user authentication in ANOTEL. They belong to the family of anonymous credentials and combine zero-knowledge proofs with efficient signature schemes. In an initial procedure, the user retrieves a so-called dispenser from the issuer. To this end, she has to unveil her identity. With the dispenser, the user can generate n anonymous tokens for each time frame and authenticate n times to a base station of the network operator. The tokens are unlinkable, and their anonymity is not revocable. If tokens are reused, though, the user's identity is automatically revealed. This way, the network operator can detect that credentials have been shared or stolen and can revoke them to prevent future misuse.

Anonymous credentials thereby protect users' privacy and respect the network operator's interests at the same time. Their use for authentication is currently undergoing an ISO standardization process [41], and the applicability for mobile devices has been shown before [42]. Therefore, they are an excellent match for ANOTEL, sharing the vision that privacy and accountability are not contradicting each other.

VII. Conclusion

With ANOTEL, we presented a novel approach to location management in cellular networks, which delivers location privacy by design. By implementing a separation of concerns between network operator, location provider, and pseudonym provider, we effectively minimize the impact of attacks on location privacy. ANOTEL hides location data from a global attacker with the capabilities of the network operator. Therefore, it invalidates the argument, that scalable cellular networks with strong location privacy are technically impossible.

At the heart of the design is a sophisticated paging protocol. Through the inclusion of encrypted caller identifiers, it becomes impossible for an attacker to silently produce signals that localize users. A dedicated channel between user and network is only established if the user initiates it. To evaluate our mechanism, we implemented a simulation scenario based on current network infrastructure statistics of Berlin, Germany. Our results show that ANOTEL successfully thwarts attackers that aim to locate and trace users. As location privacy is improved significantly even for low numbers of users and events, it is likely that the location privacy gain in real deployments is even higher.

References

[1] M. A. Bayir, M. Demirbas, and N. Eagle, "Discovering spatiotemporal mobility profiles of cellphone users," in *WoWMoM '09: Proceedings of the 10th IEEE International Symposium on a World of Wireless Mobile and Multimedia Networks*, Jun. 2009.

[2] A. R. Beresford and F. Stajano, "Location privacy in pervasive computing," in *PerCom '03: Proceedings of the 1th IEEE International Conference on Pervasive Computing and Communications*, Mar. 2003.

[3] B. Hoh and M. Gruteser, "Protecting location privacy through path confusion," in *SecureComm '05: Proceedings of the 1st IEEE/CreateNet International Conference on Security and Privacy for Emerging Areas in Communication Networks*, Sep. 2005.

[4] A. R. Beresford and F. Stajano, "Mix zones: User privacy in location-aware services," in *PerCom '04: Proceedings of the 2th IEEE International Conference on Pervasive Computing and Communications*, Mar. 2004.

[5] D. F. Kune, J. Koelndorfer, N. Hopper, and Y. Kim, "Location leaks on the GSM air interface," in *NDSS '12: Proceedings of the Network and Distributed System Security Symposium*, Feb. 2012.

[6] S. Gürses, C. Gonzalez Troncoso, and C. Diaz, "Engineering privacy by design," *Computers, Privacy & Data Protection*, 2011.

[7] J. Meyerowitz and R. Roy Choudhury, "Hiding stars with fireworks: location privacy through camouflage," in *MobiCom '09: Proceedings of the 15th Annual ACM International Conference on Mobile Computing and Networking*, Sep. 2009.

[8] C. A. Ardagna, S. Jajodia, P. Samarati, and A. Stavrou, "Providing users' anonymity in mobile hybrid networks," *ACM Transactions on Internet Technology (TOIT)*, vol. 12, no. 3, 2013.

[9] M. Gruteser and D. Grunwald, "Enhancing location privacy in wireless LAN through disposable interface identifiers: a quantitative analysis," *Mobile Networks and Applications*, vol. 10, no. 3, 2005.

[10] T. Jiang, H. J. Wang, and Y.-C. Hu, "Preserving location privacy in wireless LANs," in *MobiSys '07: Proceedings of the 5th International Conference on Mobile Systems, Applications, and Services*, Jun. 2007.

[11] H. Damker, H. Federrath, M. Reichenbach, and A. Bertsch, "Persönliches erreichbarkeitsmanagement," in *Mehrseitige Sicherheit in der Kommunikationstechnik*. Addison-Wesley-Longman, 1997.

[12] A. Pfitzmann, "Technischer datenschutz in öffentlichen funknetzen," *Datenschutz und Datensicherung*, no. 17/8, 1993.

[13] K. Sung, B. N. Levine, and M. Liberatore, "Location privacy without carrier cooperation."

[14] R. Dingledine, N. Mathewson, and P. Syverson, "Tor: The second-generation onion router," in *USENIX Security '04: Proceedings of the 13th USENIX Security Symposium*, Aug. 2004.

[15] D. Kesdogan, P. Reichl, and K. Junghärtchen, "Distributed temporary pseudonyms: A new approach for protecting location information in mobile communication networks," in *ESORICS '98: Proceedings of the 3th European Symposium on Research in Computer Security*, Sep. 1998.

[16] Q. He, D. Wu, and P. Khosla, "The quest for personal control over mobile location privacy," *IEEE Communications Magazine*, vol. 42, no. 5, 2004.

[17] C. Tang and D. O. Wu, "Mobile privacy in wireless networks-revisited," *IEEE Transactions on Wireless Communications*, vol. 7, no. 3, 2008.

[18] I. F. Akyildiz, J. S. Ho, and Y.-B. Lin, "Movement-based location update and selective paging for PCS networks," *IEEE/ACM Transactions on Networking (TON)*, vol. 4, no. 4, 1996.

[19] K. Kyamakya and K. Jobmann, "Location management in cellular networks: Classification of the most important paradigms, realistic simulation framework, and relative performance analysis," *IEEE Transactions on Vehicular Technology*, vol. 54, no. 2, 2005.

[20] H. Zang and J. C. Bolot, "Mining call and mobility data to improve paging efficiency in cellular networks," in *MobiCom '07: Proceedings of the 13th Annual ACM International Conference on Mobile Computing and Networking*, Sep. 2007.

[21] H. Federrath, A. Jerichow, D. Kesdogan, A. Pfitzmann, and D. Trossen, "Minimizing the average cost of paging on the air interface – An approach considering privacy," in *VTC '97: Proceedings of the 47th IEEE Vehicular Technology Conference*, May 1997.

[22] T. Ta and J. S. Baras, "Enhancing privacy in LTE paging system using physical layer identification," in *Data Privacy Management and Autonomous Spontaneous Security*. Springer, 2013.

[23] M. Arapinis, L. I. Mancini, E. Ritter, and M. Ryan, "Privacy through pseudonymity in mobile telephony systems," in *NDSS '14: Proceedings of the Network and Distributed System Security Symposium*, Feb. 2014.

[24] T. Engel, "SS7: Locate. Track. Manipulate." in *31th Chaos Communication Congress (31C3)*, 2014.

[25] J. Pan, S. Paul, and R. Jain, "A survey of the research on future internet architectures," *IEEE Communications Magazine*, vol. 49, no. 7, 2011.

[26] O. Goldreich, *Foundations of Cryptography: Volume 2, Basic Applications*. Cambridge University Press, 2009, vol. 2.

[27] C. Newman, A. Menon-Sen, A. Melnikov, and N. Williams, "Salted challenge response authentication mechanism (SCRAM) SASL and GSS-API mechanisms," Jul. 2010, IETF RFC 5802.

[28] A. Baraev, U. Ayesta, I. M. Verloop, D. Miorandi, and I. Chlamtac, "Technical vulnerability of the E-UTRAN paging mechanism," in *WCNC '12: Proceedings of the IEEE Wireless Communications and Networking Conference*, Apr. 2012.

[29] M. Gruteser and B. Hoh, "On the anonymity of periodic location samples," in *SPC '05: Proceedings of the 2nd International Conference on Security in Pervasive Computing*, Apr. 2005.

[30] R. Shokri, G. Theodorakopoulos, J.-Y. Le Boudec, and J.-P. Hubaux, "Quantifying location privacy," in *SP '11: Proceedings of the 32th IEEE Symposium on Security and Privacy*, May 2011.

[31] B. Wiedersheim, Z. Ma, F. Kargl, and P. Papadimitratos, "Privacy in inter-vehicular networks: Why simple pseudonym change is not enough," in *WONS '10: Proceedings of the 7th Annual Conference on Wireless On-demand Network Systems and Services*, Feb. 2010.

[32] ENAiKOON GmbH, "OpenCellID," http://opencellid.org/, 2014.

[33] T. Dittler, "Anotel simulation," https://github.com/t2d/anotel_sim, 2015.

[34] BITKOM, "63 millionen handy-besitzer in deutschland," https://goo.gl/3q7XfW, Aug. 2013.

[35] A. Smith, "Americans and text messaging," http://goo.gl/Xae04O, 2011.

[36] T. Narten, R. Draves, and S. Krishnan, "Privacy extensions for stateless address autoconfiguration in IPv6," Sep. 2007, IETF RFC 4941.

[37] D. Chaum, A. Fiat, and M. Naor, "Untraceable electronic cash," in *CRYPTO '88: Proceedings of the 8th Conference on Advances in Cryptology*, Aug. 1988.

[38] D. Chaum, "Blind signatures for untraceable payments," in *Advances in Cryptology*. Springer, 1983.

[39] 3GPP, *System Architecture Evolution (SAE); Security architecture (3GPP TD 33.401)*, ETSI, Mar. 2013.

[40] J. Camenisch, S. Hohenberger, M. Kohlweiss, A. Lysyanskaya, and M. Meyerovich, "How to win the clonewars: Efficient periodic n-times anonymous authentication," in *CCS '06: Proceedings of the 13th ACM Conference on Computer and Communications Security*, Oct. 2006.

[41] *ISO/IEC 20009:2013 Information technology – Security techniques – Anonymous entity authentication*, International Organization for Standardization, Genua, Schweiz, 2013.

[42] K. Potzmader *et al.*, "Group signatures on mobile devices: Practical experiences," in *TRUST '13: Proceedings of the 6th International Conference on Trust & Trustworthy Computing*, Jun. 2013.