

ANOTEL: Cellular Networks with Location Privacy

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—While location management is a key component of cellular networks, it is also a major privacy issue: 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. With ANOTEL we present a novel, clean-slate approach of location management in cellular networks that challenges this assumption. The design is able to route calls to users who move through cellular networks, without violating their location privacy.

I. INTRODUCTION

Mobile cell phones have become ubiquitous location tracking devices. They register to nearby base stations and expose their location, facilitating spatio-temporal mobility profiles [1].

We 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 silently update their location [2], and c) paging signals are plaintext broadcast messages and hence leak information about the callee's identity and location.

To this end, we design ANOTEL, a 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 [3]. 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. We restrict the network operator to its main task of providing the cellular infrastructure. Further, we add a location provider and a pseudonym provider as independent entities. No entity, including the network operator, shall ever be provided with both the users' identity and location information. This aim requires new, privacy-preserving signaling procedures, in particular for paging and call setup, which we devise as the key contribution of this paper.

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. Thus, ANOTEL protects from unnoticed collection of location data and enables 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. 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. An extended version of this paper can be found in [4].

We discuss related work in Sec. II. In Sec. III, we define our system and attacker model. Thereafter, we introduce ANOTEL's design and specify call establishment in Sec. IV. We present our evaluation in Sec. V. Sec. VI concludes this paper.

II. RELATED WORK

Our intention here is to hide user locations from the network infrastructure, particularly the network operator. Limiting the network operator's knowledge in cellular networks is often closely related to the specifics of GSM [5], [6]. [5] aims to host location data on user-operated devices. However, connecting to such a device constitutes a fingerprint. While the use of anonymizers, as in [6], mitigates the problem, it implies other weaknesses. In ANOTEL location data is fully decoupled from identities; neither additional user-operated components nor external anonymity services are needed.

[7] suggests to distribute user pseudonyms among n parties. An attacker may, however, still generate incoming calls and, thereby, trigger pseudonym resolution. In ANOTEL, pseudonyms are random one-time handles, and network operations, including call routing, do not need to resolve pseudonyms.

Research on paging protocols focuses mostly on efficiency [8]. [9] 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. [10] 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.

[11] criticizes the infrequent change of temporary mobile subscriber identities (TMSIs) in today's cellular networks. Location privacy, though, depends on not knowing the association between long-term identity (IMSI) and TMSI. In traditional networks, the association is always known to the network operator. Moreover, anyone with access to SS7 can resolve TMSIs [12]. ANOTEL precludes such attacks by avoiding knowledge of user locations in the first place.

III. SYSTEM AND ATTACKER MODEL

For the discussion of our protocol, we consider an abstract cellular network: it is operated by a *network operator* and clustered into *location areas (LAs)*. Each LA consists of one

or more *cells*, each covered by one *base station*. We assume that the network is much larger than a single cell's coverage area. Therefore, location information is needed to route calls.

In today's cellphone standards, the users' current locations are stored in *location registers* controlled by the network operator. When a user enters a new LA, she sends a *location update*. 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?* Therefore, we assume a semi-honest attacker with global network knowledge. 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. Such direct tracking serves as our baseline.

We do not consider attacks that require collaboration between ANOTEL's major entities. Colluding attacks would trivially subvert the separation of concerns and reduce location privacy to the level of current cellular networks.

IV. ANOTEL

The network operator N remains a central entity in our design. N manages the base stations and operates the backbone network. Therefore, N can observe all traffic. User identities, however, are never revealed to N . 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 and a temporary pseudonym. Only pseudonyms are visible to N . While P knows the mapping between pseudonyms and user IDs, P never gains any information about user locations.

The current user locations are stored at a third entity, the *location provider* L . L maintains a mapping from pseudonyms to locations but is not able to determine user identities.

The separation between N , P , and L is a core feature of ANOTEL's design. Below, we describe each component's role in more detail, focusing on ANOTEL's signaling protocol as a key ingredient of user privacy.

A. Key and Pseudonym Management

We assume P shares symmetric keys $K_{P,N}$ and $K_{P,L}$ with N and L , respectively. Furthermore, P and L each have asymmetric encryption key pairs; their public keys K_P and K_L are assumed to be known by all users. $\{m\}_{K_{x,y}}$ denotes that a message m is encrypted using a symmetric key $K_{x,y}$ shared between x and y . $\{m\}_{K_x}$ denotes asymmetric encryption of m using x 's public key.

Users have a long-term identity id , comparable to an IMSI, and a temporary pseudonym p . Pseudonym changes are performed frequently: upon every location update, outgoing call, and time-based. To update p , ANOTEL user Alice (A) chooses a new pseudonym at random. Then, A transmits p to P in a *pseudonym update* message $A \rightarrow P: \{id, p, k, t\}_{K_P}$.

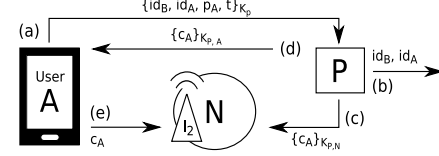


Fig. 1. Outgoing call signaling (i.e., from ANOTEL to a foreign network).

The timestamp t ensures freshness; k authenticates A towards P and may be replaced by some other form of authentication, like [13], in practice.

Like any cellular network, ANOTEL needs to keep track of user locations. This task is outsourced to L . Users send *location updates* to L when changing LAs. A 's update for example looks like $A \rightarrow L: l, \{p, s\}_{K_L}$. It contains A 's pseudonym p , the current LA l , and a random seed s , which will be used for paging as discussed below. A knows her LA from periodic beacons sent by the base stations. p and s must not be readable by N . l , in contrast, is used for a sanity check by N (who knows l anyway). Note, L cannot map pseudonyms to users. For this reason, L is neither able to correlate subsequent pseudonyms of the same user nor to observe the duration for which a pseudonym remains in a specific location.

Location updates add a new record only and do not remove or refresh a previous pseudonym. The location information at L time out after a certain time span. To harden against exploits, time-based updates are performed randomly within the second half of the location record's lifetime.

B. Call Establishment

Call setup, including routing and paging, 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.

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

Fig. 1 shows our protocol step by step. First, A signals P her intention to call B (Fig. 1(a)). The message includes her pseudonym p_A and a timestamp t . P acts as a signaling proxy for users, which can issue and receive calls without N being able to see the involved ids .

After receiving A 's message, P checks t for freshness and validates id_A and p_A . If checks succeed, P signals A 's call to the foreign network (Fig. 1(b)). Inter-network signaling remains unchanged. If signaling succeeds, P sends an encrypted anonymous token c_A to N and A : $P \rightarrow N: \{c_A\}_{K_{P,N}}$, $P \rightarrow A: \{c_A\}_{K_{P,A}}$. Finally, A reveals c_A to N (Fig. 1(e)), which authenticates A for the initiated call.

In contrast to traditional implementations, N neither learned that A and B are talking nor where—only that “some user” in l_2 has established a call. P knows A 's identity but not

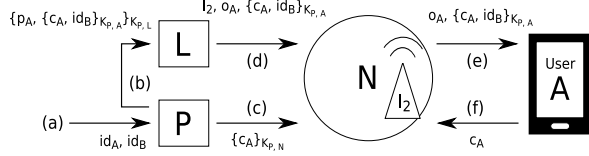


Fig. 2. Incoming call signaling (i. e., from a foreign network to ANOTEL).

her current location. Both N and P are thus unable to create individual movement profiles.

The timestamp in A 's first message prevents replay attacks. It 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 A 's identity without knowledge of c_A . Once a token has been revealed in step (e), it cannot be reused.

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

First, the foreign network signals B 's call request to P (Fig. 2(a)). P translates id_A to A 's current pseudonym p_A and waits for the next paging cycle, while in the meantime collecting further incoming call requests.

The waiting time mixes incoming calls and prevents L from linking pseudonyms to users. A paging cycle should mix several calls without impacting user experience; thus, mixing is easier for networks with high call frequencies.

In the next paging cycle, P sends a message to L to initiate paging (Fig. 2(b)). The message structure for B 's call request is $\{p_A, \{c_A, id_B\}_{K_{P,A}}\}_{K_{P,L}}$, where c_A is an 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.

The pseudonym provider also informs N about B 's call (Fig. 2(c)). Recall that during the location update, A transmitted a random seed s to L . 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 A can compute paging names. In particular, A 's paging names are unlinkable. At the same time, the hash-based generation of o_i is computationally efficient for both L and A .

Paging names require synchronization between L and A . Since there is no transmission reliability for pagers, A may miss pagers. Therefore, A listens 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 A 's current paging name o_A , L constructs the paging signal for l_2 (Fig. 2(d)). This process is repeated for every incoming call in P 's list. For calls in the same location area, the paging information is appended. The result is a list of pagers with one packet per LA. These paging packets are subsequently padded with dummies, i. e., paging messages

to random, fake paging names. Dummies prevent observers from building usage profiles based on the existence or size of packets for different LAs.

L hands over the paging packets (one per LA) to N . For A 's location, the paging packet looks like: $l_2, o_A, \{c_A, id_B\}_{K_{P,A}}, \dots, o_m, \{c_m, id_{caller_m}\}_{K_{P,U_m}}$. Here, m denotes the maximum number of pagers per LA and paging cycle. For example, $m = 56$ pagers in a 1280 ms period for LTE [14]. The overhead incurred by padding is negligible in relation to a cell's capacity. 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. Likewise, an eavesdropper cannot distinguish pagers. Without knowledge of A 's current paging name, the pagers do not reveal her current location.

N broadcasts the paging packets (Fig. 2(e)). A 's phone listens for pagers and checks if any 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, her phone rings to inform A about the call from B . In contrast to traditional cellular networks, no session between N and A has yet been established.

Only if A decides to answer the call, A reveals c_A to N (Fig. 2(f)). In this case, N knows that someone is calling a user located in l_2 , but N knows neither the identity of the caller nor that of the callee. P knows that A and B 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 ANOTEL calls, are implemented using a combination of incoming and outgoing signaling, paired via P .

In summary, the major distinctive feature of ANOTEL is the separation between N , L , and P . We have shown how our system design uses these entities to decouple identities from location management and network operation.

V. EVALUATION

We model the cellular network using a grid of cells, which are randomly grouped into rectangular LAs. Our scenario is supposed to resemble a rush hour in Berlin, Germany. We use the OpenCellID data [15] from 2014 to parametrize the grid construction. On average, each of the major network operators manages around 2,500 UMTS cells and 42 LAs in Berlin [16].

We implemented a discrete event simulation, which simulates user movement and call events. During the simulation, users move between randomly chosen pairs of coordinates at constant speed. We choose users' individual velocities such that their journeys take the whole simulation time of 2 h. Calls are randomly and uniformly distributed over this time. We assume the initial user location to be known to the attacker.

ANOTEL is designed to protect users against a global attacker. We measure the privacy gain using the reduction of tracking success, which is a common privacy metric [17], [18]. In today's networks, attackers can trivially trace users by listening to location updates, paging, and call establishment. Thus, the tracking success of the attacker will always be 100%.

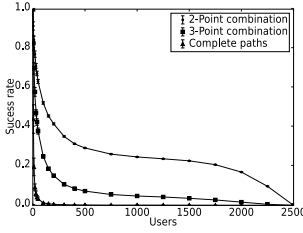


Fig. 3. Impact of the number of users on the success rate (6 events, 50 runs, 95 % confidence intervals).

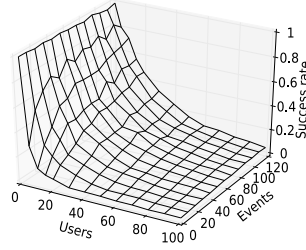


Fig. 4. Impact of the number of users and events on the success rate (complete paths, mean of 50 runs).

In ANOTEL, an attacker is not capable of linking consecutive pseudonyms of the same user. Any attempt to heuristically track user movements results in ambiguous mappings. This effect is comparable to mix zone approaches [19], 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.

To quantify ANOTEL's privacy gains, we implement a tracking heuristic: whenever a new signal occurs, the attacker assigns it to the assumed user location closest to the location of the new signal. Our heuristic, together with the scenario and evaluation methods, is publicly available [16].

Fig. 3 illustrates the impact of the number of users on the fraction of correctly matched tracks. We call the latter the attacker's success rate. This success rate is calculated for sub-paths of length two and three and for complete paths. If there is only one user per LA, the attacker can obviously track her with certainty. With an increasing number of users, the success rate drops rapidly, due to the above-mentioned mixing effect. The remaining tracking success is mainly due to users located in sparsely populated areas. The attacker's tracking heuristic completely fails when there are more users than cells.

Tracking success is also influenced by the number of events. It is easier to trace users if they communicate more frequently. Fig. 4, 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).

We further investigate the expected number of pagings. In its inner city, Berlin has about 1.7 million inhabitants. As 90 % of Germans own a cellphone [20], there are 612 phones per cell on average. According to [21], an average American adult sends or receives 42 messages and 12 calls per day. Numbers for Germany are not available, but likely comparable. As every event is counted twice (incoming and outgoing), we get 27 pagings per day and person. Thus, an attacker would likely see more than ten pagings per second and LA:

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

The tracking heuristic, as discussed before, is likely to fail under these conditions.

In summary, ANOTEL's mixing of pseudonyms hinders an attacker from tracking all users directly. 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. 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.

REFERENCES

- [1] M. A. Bayir, M. Demirbas, and N. Eagle, "Discovering spatiotemporal mobility profiles of cellphone users," in *WoWMoM '09*, Jun. 2009.
- [2] D. F. Kune, J. Koelndorfer, N. Hopper, and Y. Kim, "Location leaks on the GSM air interface," in *NDSS '12*, Feb. 2012.
- [3] S. Gürses, C. Gonzalez Troncoso, and C. Diaz, "Engineering privacy by design," *Computers, Privacy & Data Protection*, 2011.
- [4] T. Dittler, F. Tschorsch, S. Dietzel, and B. Scheuermann, "ANOTEL: Cellular networks with location privacy (extended version)," *Cryptology ePrint Archive*, Report 2016/764, 2016, <https://ia.cr/2016/764>.
- [5] A. Pfizmann, "Technischer datenschutz in öffentlichen funknetzen," *Datenschutz und Datensicherung*, no. 17/8, 1993.
- [6] K. Sung, B. N. Levine, and M. Liberatore, "Location privacy without carrier cooperation."
- [7] D. Kesdogan, P. Reichl, and K. Junghärtchen, "Distributed temporary pseudonyms: A new approach for protecting location information in mobile communication networks," in *ESORICS '98*, Sep. 1998.
- [8] 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.
- [9] H. Federrath, A. Jerichow, D. Kesdogan, A. Pfizmann, and D. Trossen, "Minimizing the average cost of paging on the air interface – An approach considering privacy," in *VTC '97*, May 1997.
- [10] 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.
- [11] M. Arapinis, L. I. Mancini, E. Ritter, and M. Ryan, "Privacy through pseudonymity in mobile telephony systems," in *NDSS '14*, Feb. 2014.
- [12] T. Engel, "SS7: Locate. Track. Manipulate." in *31th Chaos Communication Congress (31C3)*, 2014.
- [13] 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.
- [14] A. Baraev, U. Ayesta, I. M. Verloop, D. Miorandi, and I. Chlamtac, "Technical vulnerability of the E-UTRAN paging mechanism," in *WCNC '12*, Apr. 2012.
- [15] ENAiK00N GmbH, "OpenCellID," <http://opencellid.org/>, 2014.
- [16] T. Dittler, "Anotel simulation," https://github.com/t2d/anotel_sim, 2015.
- [17] M. Gruteser and B. Hoh, "On the anonymity of periodic location samples," in *SPC '05*, Apr. 2005.
- [18] R. Shokri, G. Theodorakopoulos, J.-Y. Le Boudec, and J.-P. Hubaux, "Quantifying location privacy," in *SP '11*, May 2011.
- [19] A. R. Beresford and F. Stajano, "Location privacy in pervasive computing," in *PerCom '03*, Mar. 2003.
- [20] BITKOM, "63 millionen handy-besitzer in deutschland," <https://goo.gl/3q7XfW>, Aug. 2013.
- [21] A. Smith, "Americans and text messaging," <http://goo.gl/Xae04O>, 2011.