IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received December 25, 2017, accepted January 31, 2018, date of publication February 28, 2018, date of current version April 23, 2018.

Digital Object Identifier 10.1109/ACCESS.2018.2808333

Secure and Fault-Tolerant Distributed Location Management for Intelligent 5G Wireless Networks

KASHIF MUNIR[1], EHTESHAM ZAHOOR[1], RAFIA RAHIM[1], XAVIER LAGRANGE[2], (Senior Member, IEEE), AND JONG-HYOUK LEE[3], (Senior Member, IEEE)

[1]National University of Computer and Emerging Sciences, Islamabad 46000, Pakistan
[2]IMT Atlantique, IRISA, 35576 Cesson-Sévigné, France
[3]Department of Software, Sangmyung University, Cheonan 330-720, South Korea

Corresponding author: Jong-Hyouk Lee (jonghyouk@smu.ac.kr)

This work was supported by the Cross-Ministry Giga KOREA Project grant funded by the Korean Government (MSIT) (No. GK18P0400, Development of Mobile Edge Computing Platform Technology for URLLC Services).

ABSTRACT Distributed mobility management (DMM) requires a flattened cellular network architecture to meet ever-increasing traffic demands of mobile users. In order to get rid of centralized mobility entities, it is desirable to manage locations of mobile nodes (MNs) in a distributed but secure and fault-tolerant manner. To address such issues, in this paper, we propose a secure and fault-tolerant mechanism for DMM. The proposed mechanism utilizes a distributed hash table of access nodes for distributed location management, and a ticket-reuse approach is adopted for improving authentication performance while providing secure authentication of the MNs. We present an analytical model to calculate the processing time of a location query. We have evaluated the proposed security mechanism in terms of authentication latency and a number of authentication messages required for secure handover. Analysis results confirm that the proposed mechanism outperforms the popular existing authentication protocol, EAP-TLS, in a DMM environment.

INDEX TERMS Handover, secure authentication, 5G.

I. INTRODUCTION

Existing mobility management protocols such as Mobile IPv6 (MIPv6) [1] and Proxy Mobile IPv6 (PMIPv6) [2] have been developed with an assumption that a cellular network architecture is centralized. However, the assumption is being changed. The cellular network architecture is being flattened for better performance and scalability, while there are many reports that centralized network architecture and mobility management protocols cause a network bottleneck and present a single point of failure. To address the centralized mobility management protocol issues, the IETF Distributed Mobility Management (DMM) working group was proposed to develop DMM protocols to operate in a distributed network architecture [3] for intelligent 5G wireless networks [4]–[8]. Some of the important ongoing works of the working group include DMM deployment models and architectural considerations, on-demand mobility management, protocol for forward policy configuration, etc.

In DMM, the data plane and the control plane are separated and the data plane is distributed for better performance and scalability over the network architecture. The control plane however can either be distributed (i.e., fully distributed DMM) or centralized (i.e., semi-distributed DMM) for ease of location management [9], [10]. For instance, in [11], a centralized Session Initiation Protocol (SIP) register has been proposed for a DMM environment. However, the SIP register can be duplicated to improve the reliability but it does not give the level of resilience that can be obtained with a fully distributed architecture. There is thus a need to define an architecture fully compatible with DMM. More details about the background and approaches of DMM are available in [9] and [10]. In addition, as a preliminary work [12], we analyzed the performance of a DMM approach called Dynamic Mobility Anchoring (DMA) [13] in comparison with PMIPv6.

Without securing wireless network access, an illegitimate Mobile Node (MN) can access network resources and launch various attacks. Only an authenticated and authorized MN, i.e., legitimate MN, should be able to access the network resources. For instance, only a legitimated MN registers

and updates its location during its handover after secure authentication. If no wireless network access security provided, an illegitimate MN can send malicious location update messages on behalf of a legitimate MN or it can redirect or block data packets related to location updates. To address such issues, in this paper, we propose a secure and fault-tolerant mechanism that uses a distributed hash table of access nodes for distributed location management while a ticket-reuse approach is adopted for improving authentication performance. The main contributions of this paper are as follows:

1) Fault-tolerant and fully distributed location management based on a Chord Distributed Hash Table (DHT).
2) Analytical model to calculate the average processing (response) time of a location query.
3) Ticket-reuse approach designed for distributed location management.
4) Performance analysis results of the proposal that has been compared with TAP-TLS in terms of authentication latency and number of authentication messages.

The remainder of the paper is organized as follows: Section II briefly describes the literature closely related to our work. Section III presents the analytical model to calculate the processing time of a location query and describes the proposed mechanism. Performance analysis is described in Section IV and the results are presented in Section V. The paper is concluded with some remarks in Section VI.

II. RELATED WORK

A. MOBILITY MANAGEMENT AND HANDOVER AUTHENTICATION

A distributed and dynamic location management scheme proposed in [14] reduces the signaling traffic and signaling delay as compared to previously developed schemes. The authors proposed the scheme to distribute the signaling burden by dynamically adjusting regional network boundaries depending on the up-to-date mobility and traffic load for each MN.

The idea of using multiple location servers instead of a centralized home location register is not new. Prakash *et al.* [15] proposed a distributed location management strategy that is based on quorums and dynamic hashing and has the granularity of a Registration Area (RA). The works presented in [16]–[19] have similar ideas. The main aim of the mechanism in [15] was to provide load balancing among location servers in order to reduce location update and query time.

A dynamic location management scheme is described in [20] for the low mobility rate of an MN and a static scheme is presented for the high mobility rate. According to the authors' claim, the dynamic scheme performs better than the static scheme for high paging cost or when the number of cells in an Location Area (LA) is large.

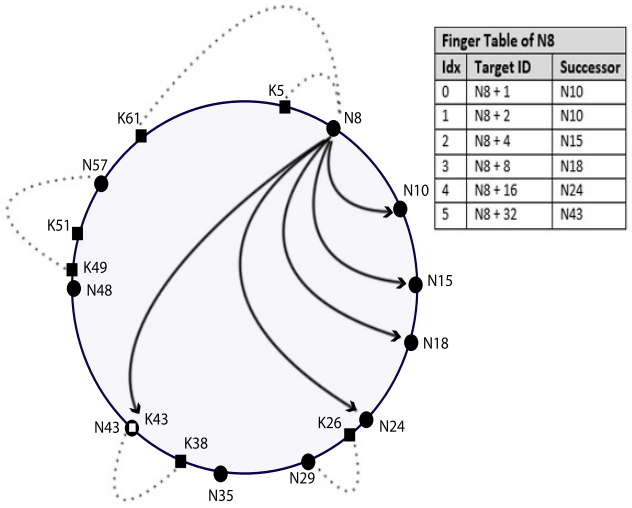
Sidhu and Singh [21] describe several metrics such as location update cost and paging cost metrics for location management. The schemes are classified into static and dynamic location update schemes. The static schemes are based on the topology of the network whereas the dynamic schemes are based on time, movement or distance. In [22], a simulation environment is described with a claim that, in terms of location management cost, LA based location management using a profile or history-based direction information gives the best performance. In order to reduce the cost of location updates, a dynamic location management approach has been proposed in [23]. However, the approach slightly increases the paging cost compared to the static approach.

DHT based mechanisms have extensively been proposed to solve different problems in wireless networks. An improved handover management mechanism based on a DHT was proposed in [24]. A DHT based distributed localization service for wireless mesh networks is also proposed in [25]. To reduce the packet loss due to IP handovers, a mechanism utilizing the mobile Stream Control Transmission Protocol (mSCTP) for improved transport and a Chord DHT [26] for location management was proposed in [27]. A two-tier Chord DHT for distributed location management was introduced in [28] that uses the nodes with high computation power and stability as location servers.

For DMM, a SIP-based centralized location server was described in an earlier work of Ali-Ahmad *et al.* [29]. The main objective of the work in [29] was to reduce the number of location update messages that are sent by MNs to the centralized location server. An analytical model was presented to calculate the load of the location updates.

TAP-TLS [30] is a popular authentication scheme that is widely used in existing wireless mobile networks. It performs the full EAP exchange for initial and handover authentications for an MN. As a baseline authentication scheme, TAP-TLS has been deployed and analyzed, but as studied, this scheme does not provide any performance optimization. For instance, when an MN changes its attachment point, the MN must perform the full EAP exchange that causes a long latency so that user's quality of experience is degraded. To address this kind of problem, studies on performance optimization in handover authentication have been conducted. For instance, the authentication scheme in [31] utilizes a local credential obtained from an Access Router (AR), i.e., Mobile Access Gateway (MAG) of PMIPv6, in initial authentication when an MN performs handover authentication with another AR in a PMIPv6 environment. In [35], Handover Optimized Ticket-based Authentication (HOTA) has been proposed in which an MN performs handover authentication on different access networks securely by reusing credential issued by authentication server in a PMIPv6 environment.

The proposed work is different compared with the distributed and dynamic location management scheme in [14] as the proposed mechanism considers a fixed-size access network area, i.e., the area covered by the AR. In addition, the proposed mechanism introduces a replication strategy for a distributed and secure location management based on a DHT. Our work is also different compared with the previous DHT based mechanisms [24], [25], [27], [28] as the proposed mechanism considers that all ARs have location



Finger Table of N8		
Idx	Target ID	Successor
0	N8 + 1	N10
1	N8 + 2	N10
2	N8 + 4	N15
3	N8 + 8	N18
4	N8 + 16	N24
5	N8 + 32	N43

FIGURE 1. A 6-bit Chord identifier space in which the dotted lines indicate which nodes host which keys and the black lines represent the fingers of node *N*8 [36].

registers and extends the DHT to provide fault tolerance in location management. Compared with the existing authentication schemes in [31] and [35], the proposed one is developed for DMM and focuses on providing an application level security.

B. CHORD DHT

Chord is a protocol and algorithm for a peer-to-peer DHT. The central pillars of the DHT are n-bit identifiers in the range $[0, 2^n - 1]$. An identifier of a node is its ID. An identifier of a data item is its key (K). A data item has a record value (V) that is associated with the key (K) of the data item. Both, the node and the data items, are mapped on the same n-bit Chord identifier space. The key and value pair (K-V) is hosted by the node whose ID is greater than or equal to K. A node in a Chord circle is responsible for all keys that precede it counter-clockwise till the previous node [36].

Figure 1 illustrates a 6-bit Chord identifier circle having ten nodes and seven data items. It shows that $N8$ is the successor node of $K5$ as it is the next node to it clockwise and the successor node of $K43$ is $N43$ as their identifiers are equal.

Every node maintains a routing table called a finger table pointing to some other nodes on the identifier circle. The finger table has n entries in a circle with n-bit identifiers. On a node p, the finger table entry at row i identifies the first node that succeeds p by at least 2^{i-1}, i.e., successor $\left(p + 2^{i-1}\right)$, where $1 \leq i \leq n$. For example, the third finger $(8 + 2^2 = 12)$ is $N15$. The i-th finger of a node in its finger table is always its immediate successor node on the identifier circle. For routing purpose, each finger entry consists of a node ID, its IP address, and port number. When a query reaches a node p such that K lies between p and the numerically closest successor in the finger table of p, then the node p reports the successor as the answer to the query.

III. SECURE AND FAULT-TOLERANT MECHANISM

In this section, we present the proposed secure and fault-tolerant mechanism. We choose a finger-table based Chord DHT to implement fault-tolerant distributed location management. Chord is preferred because it is a simple and efficient protocol for storing, updating, and retrieving a location binding of an MN in $O(logm)$ time, where m is the total number of access networks in a cellular network.

In the proposed mechanism, every MN is uniquely identified by its ID information, e.g., SIP-like URI such as *SIP:MN0@abc.com* or permanent IP address of the MN. The mobile network is made of Access Nodes (ANs). In a centralized mobile network, the Location Register (LR) is a standalone entity. We propose to distribute the LR function on every AN. It means that an AN has both AR and LR functions. MNs can access the Internet through the AR of an AN. An IP prefix is associated to that AR. In other words, when an MN is managed by an AN j, it gets an IP address whose prefix is the one associated with the AR of AN j. Location of an MN is managed by an LR, which knows the association between the ID and the prefix of an MN. So, knowing the location of an MN is equivalent to knowing its IP address in DMA for instance [13].

The data items are the IDs of the MNs and the nodes are the ANs. We map the keys and ANs on a Chord circle using hash values produced from the IDs of MN and network prefixes of ANs. Each MN has a unique ID, ID_{MN}, which is used to produce K. The hash values are mapped on the same address space. The LR maintains the location binding information of MNs. Each entry of LR holds a K-V pair. In the proposed scenario, K is the hash of ID_{MN} and V is the current IP address of the MN.

A. ADDITION/UPDATE OF K-V PAIR

We propose a fault-tolerant distributed location server. We achieve the fault-tolerance through a single level of replication for the location bindings of MNs. The chances that both peers (i.e., ANs), maintaining the main and the backup copies of an MN's location binding, go down simultaneously are close to zero. So, it is reasonable to consider only a single level of replication.

Figure 2 shows a Chord circle having 8 ANs mapped on it. It shows that the backup copy of a K-V pair, maintained on $AN7$, is accessed by a Correspondent Node (CN) in case of failure of the main copy, maintained on $AN3$. We outline the step-wise procedures of addition/update of main and backup copies of a K-V pair, separately.

1) ADDITION/UPDATE OF MAIN COPY

On addition/update of the main copy of K-V pair, the pair is assigned to the peer that is the first successor of $K = hash(ID_{MN})$ on the Chord circle.

Through the currently attached AN, an MN sends the location registration/update to its default LR. The location information that is stored along with the key is the current

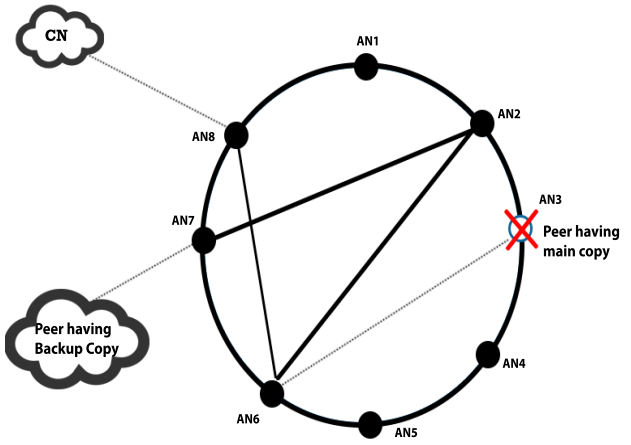


FIGURE 2. Fault-tolerance in the distributed location management. The backup copy of *K-V* pair is accessed in case of failure of the main copy.

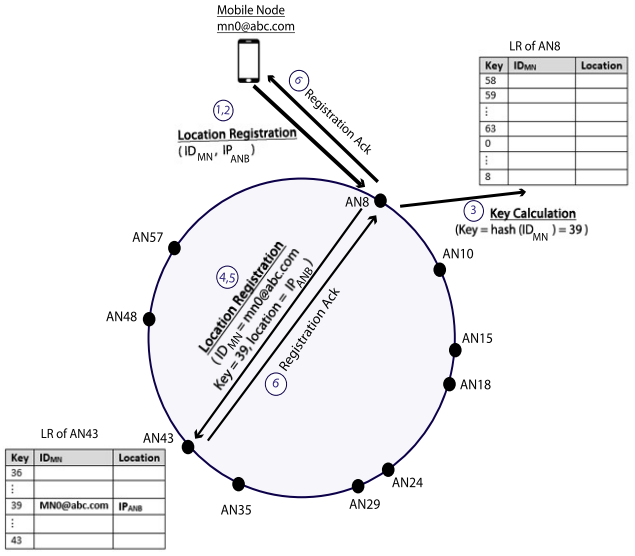


FIGURE 3. Addition/update of main copy of *K-V* pair.

IP address of the MN. We call this query as a Location Update Query (LUQ).

When a peer receives a location registration/update query from an MN, it checks whether the address space managed by it includes K. In case the address space includes K, it registers/updates the location and sends the acknowledgement to the MN. In case, it does not include K, it checks its finger table to find the closest successor of the K on the Chord circle and then it forwards the query to that successor node. This process continues till the target node, managing the relevant K, is found out. With the help of a simple example, in which 10 ANs are mapped on a 6-bit Chord circle, we explain the *K-V* pair addition/update process of the proposed scheme. Figure 3 illustrates the *K-V* pair addition/update process.

1) An MN is connected with its AN (i.e., $AN8$ in Figure 3) to register/update its location.
2) $AN8$ sends MN's location registration/update request to its LR.
3) The LR of $AN8$ calculates $K = hash(ID_{MN})$. The LR of $AN8$ checks whether the address space managed by it includes K. The LR does not find K. So, $AN8$ checks its finger-table in order to find the closest successor of K on the Chord circle. As shown in Figure 3, the closest successor to K is $AN43$.
4) $AN8$ forwards the locations registration/update query to $AN43$.
5) The LR of $AN43$ finds out that K is managed by it. So, it stores/updates the location binding information of the MN.
6) Subsequently, the LR of $AN43$ replies back the registration acknowledgement on the reverse query path to the MN.

The Location Search Query (LSQ) from a CN is handled in the same way as mentioned above. As a result of the LSQ, the CN gets the current IP address of the MN on providing the ID_{MN}.

2) ADDITION/UPDATE OF THE BACKUP COPY

On addition/update of the backup copy of *K-V* pair, the pair is assigned to the peer which is the first successor of $K' = Successor(hash(ID_{MN})+2^n/2)$ on the Chord circle. The steps of addition/update of the backup copy of *K-V* pair are the same as explained above for the addition/update of the main copy. We emphasize that the failure of an AN in the proposed mechanism is a rare event as opposed to a typical P2P system.

B. ANALYTICAL MODEL FOR CALCULATING THE PROCESSING TIME

In this section, we present an analytical model for calculating the average processing time taken by a Location Query (LQ). An LQ can either be a Location Update Query (LUQ) or a Location Search Query (LSQ). Since, each type of query performs the same type of routing steps in the DHT and undergoes similar type of processing so, in the model, we do not differentiate them and consider each of them as an LQ. An LQ can either have a query-miss on an AN or it can have a query-hit on the AN. In case of query-miss, the AN is called an intermediate AN and in case of query-hit, it is called a final AN. The two types of processing at an AN results in a two-queue based query processing system as shown in Figure 4. Each queue of the system is a deterministic queue having its own service rate. Since the processing steps at an AN are always same, an LQ always takes a constant service time for processing while being in service. The symbols that are used in the analytical model are listed in Table 1.

We assume that LQ requests arrive at an AN according to a Poisson process. The process results naturally when a reasonably large population of the MNs update their locations independently and a reasonably large population of the CNs search for the locations of the MNs independently. Since there is no practical evidence that the pattern of arrivals of LQ requests on an AN follows any special distribution, but there are strong indications that a cellular network have a large numbers of MNs, the Poisson assumption holds.

TABLE 1. List of the symbols used in the analytical model.

Symbol	Meaning
λ_{LSQ}	Arrival rate of LSQs on the DHT related to an MN
λ_{LUQ}	Arrival rate of LUQs on the DHT sent by an MN
λ_{LQ}	Arrival rate of both types of LQs on the DHT
λ	Arrival rate of LQs on an AN
λ_i	Arrival rate of the LQs for which an AN is intermediate
λ_f	Arrival rate of the LQs for which an AN is final
μ_i	Service rate of an LQ on an intermediate AN
μ_f	Service rate of an LQ on a final AN
ρ_i	Utilization of the server of the $m/D_i/1$ queue
ρ_f	Utilization of the server of the $m/D_f/1$ queue
W_i	Waiting (processing) time of an LQ on the intermediate AN
W_f	Waiting (processing) time of an LQ on the final AN
N_{MN}	Total number of MNs
N_{AN}	Total number of ANs
N_{DHT}	Average number of DHT overlay hops that are traversed for an LQ
N_{hops}	Average number of the physical IP network routing hops to be traversed by an LQ for each hop of the DHT
R_{delay}	Average per hop routing delay of the IP network
Θ	Average processing time of an LQ

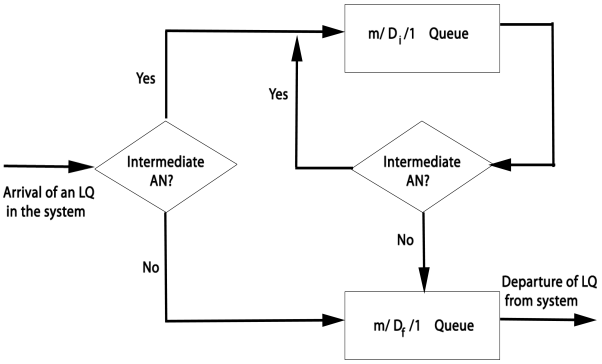


FIGURE 4. The two-queue based query processing system wherein D_i and D_f represent the deterministic service times of an LQ at intermediate and final ANs, respectively.

Let λ_{LSQ} be the arrival rate of LSQs on the DHT related to an MN. This is the rate with which the CNs are creating the sessions on the MN. Let λ_{LUQ} be the arrival rate of LUQs on the DHT sent by an MN. This is the rate with which the MN is updating its location in the DHT. Let λ_{LQ} be the arrival rate of both types of LQs on the DHT. Then, λ_{LQ} can be calculated as follows:

$$\lambda_{LQ} = \lambda_{LSQ} + \lambda_{LUQ} \tag{1}$$

Let N_{MN} be the total number of MNs and N_{AN} be the total number of ANs in the cellular network. Let λ be the arrival rate of LQs on an AN. Then, λ is calculated as follows:

$$\lambda = \frac{\lambda_{LQ} N_{MN}}{N_{AN}} \tag{2}$$

Let N_{DHT} be the average number of DHT overlay hops that are traversed for an LQ. Based on the findings in [32]–[34], it can be calculated as follows:

$$N_{DHT} = \left\lceil \frac{N_{AN}}{\log(\log(N_{AN}))} \right\rceil \tag{3}$$

Let p be the probability that the arrival of an LQ is on an intermediate AN. Then, $p = \frac{N_{DHT}}{N_{DHT}+1}$. Let $\lambda_i (= p\lambda)$ be the arrival rate of the LQs for which an AN is an intermediate AN. Let μ_i be the service rate of an LQ on an intermediate AN and $\rho_i \left(= \frac{\lambda_i}{\mu_i}\right)$ be the utilization of the server of this $m/D_i/1$ queue. Let W_i be the waiting (processing) time of an LQ on the intermediate AN. Then, from the standard queueing theory, W_i is calculated as follows:

$$W_i = \frac{1}{\mu_i} + \frac{\rho_i}{2\mu_i(1-\rho_i)} \tag{4}$$

Let $\lambda_f (= (1-p)\lambda)$ be the arrival rate of the LQs for which an AN is a final AN. Let μ_f be the service rate of an LQ on a final AN and $\rho_f \left(= \frac{\lambda_f}{\mu_f}\right)$ be the utilization of the server of this $m/D_f/1$ queue. Let W_f be the waiting (processing) time of an LQ on the final AN. Then, from the standard queueing theory, W_f is calculated as follows:

$$W_f = \frac{1}{\mu_f} + \frac{\rho_f}{2\mu_f(1-\rho_f)} \tag{5}$$

Let N_{hops} be the average number of the physical IP network routing hops to be traversed by an LQ for each hop of the DHT and R_{delay} be the average per hop routing delay of the IP network. Let Θ be the average processing time of an LQ. Then, Θ can be calculated as follows:

$$\Theta = (N_{DHT} \times N_{hops} \times R_{delay}) + (N_{DHT} \times W_i) + W_f \tag{6}$$

C. ADDING SECURITY

We present Ticket-based Authentication for Location Management (TALM) of the proposal. The authentication process can be divided into two parts. As explained in Section III, we define $AN = AR + LR$, i.e., an AN has both AR and LR functions.

1) Authentication at the time of expiry of the required security credentials (for performing authentication) of

TABLE 2. List of the symbols used in TALM.

Symbol	Meaning
$A\|\|B$	Concatenation of messages A and B
K_{A-B}	Key shared between A and B
$E(K,X)$	Message X encrypted using key K
N_A	Nonce (random number) generated by A
TK	Ticket
TT	Start time and End time

an MN or at the first time the MN is switched on (Ticket creation and expiry).

2) Authentication at the time of an IP handover of an MN in which the MN has to obtain the required security credentials from the previous AR (Ticket collection from the previous AR).

The symbols that are used in TALM are listed in Table 2. We assume that the keys shared between the Authentication Server (AS) and the MN, K_{AS-MN}, and between the AR and the AS, K_{AS-AR}, already exist. The AS only needs to provide K_{MN-AR} (i.e., session key), which will be shared between the MN and the AR.

1) TICKET CREATION AND EXPIRY

Following are the steps of TALM for ticket creation and expiry.

1) An initial message (MN_{Reg}) is sent from the MN to the AR when the MN has just been switched on and needs to register itself with the AR or MN is creating a new ticket due to the expiry of the previous ticket.

$$MN->AR : MN_{Reg}(ID_{MN}||ID_{AR}||N_{MN})$$

where N_{MN} is the nonce generated by the MN.

2) As the AR receives this initial message, it adds its own nonce and creates a message ($MN_{Reg}AS_{Msg}$) to be sent to the AS:

$$AR->AS : MN_{Reg}AS_{Msg}(E(K_{AR-AS}, ID_{MN}||ID_{AR}||N_{MN}||N_{AR}))$$

The AR encrypts this message before sending to the AS using the key it shares with the AS.

3) The AS generates the following ticket:

$$TK = E(K_{TK}, K_{MN-AR}||ID_{MN}||TT)$$

This ticket is encrypted using another key K_{TK}. Inside this ticket is K_{MN-AR} (that is shared between the MN and AR) and the time TT for which this ticket is valid. The AS wants to transmit the K_{MN-AR} to both the MN and the AR. Since, sending important data such as a key over the network is not safe, we use K_{TK}.

4) Now the AS will send a message (AS_{Res}) to send K_{MN-AR} to the AR by encrypting it using the key it shares with the AR:

$$AS->AR : AS_{Res}(ID_{MN}||TK||E(K_{MN-AS}, K_{MN-AR}||ID_{AR}||TT||N_{MN})||E(K_{AR-AS}, K_{TK}||TT||N_{AR}))$$

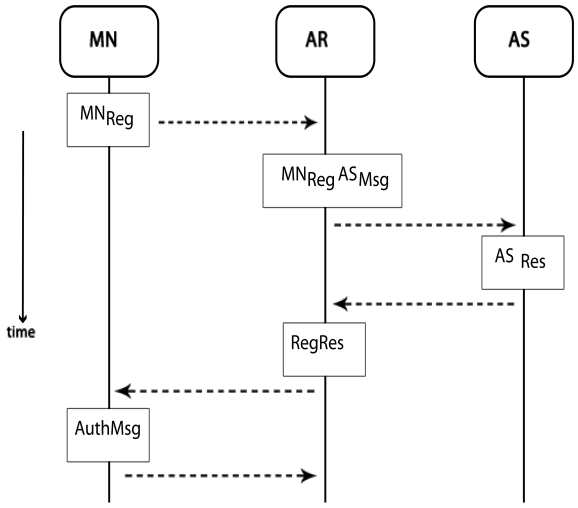


FIGURE 5. Ticket creation and expiry.

The AR will only be able to extract

$$E(K_{AR-AS}, K_{TK}||TT||N_{AR})$$

part of the message. Once the AR gets K_{TK}, it can decrypt TK and extract K_{MN-AR} (the key it shares with the MN).

5) Now the AR will communicate K to the MN using ($RegRes$) message:

$$AR->MN : RegRes(ID_{MN}||TK||E(K_{MN-AS}, K_{MN-AR}||ID_{AR}||TT||N_{MN}))$$

Using its key with the AS, the MN will be able to extract K_{MN-AR}.

6) Now both AR and MN have K_{MN-AR} and both have knowledge of the ticket's expiry. A last message ($AuthMsg$) will be sent by the MN to the AR to authenticate itself along with a location update message.

$$MN->AR : AuthMsg(ID_{AR}||TK||E(K_{MN-AR}, ID_{MN}||IP-Addr||(N+1)_{MN}))$$

If the AR successfully decrypts this message, then the MN has successfully authenticated itself. $IP\text{-}Addr$ is the current IP address of the MN, which is securely sent. Nonces are used in the messages to prevent man-in-the-middle replay attack. The flow of the messages is shown in Figure 5.

2) TICKET COLLECTION FROM A PREVIOUS AR

Following are the steps of TALM for ticket collection from previous AR.

1) The MN sends a message (nAR_{Reg}) to the nAR (newly attached AR of the MN), which does not know about K_{MN-AR}.

$$MN->nAR : nAR_{Reg}(TK||ID_{MN}||N_{MN}+1)$$

where N_{MN} is a nonce generated by the MN.

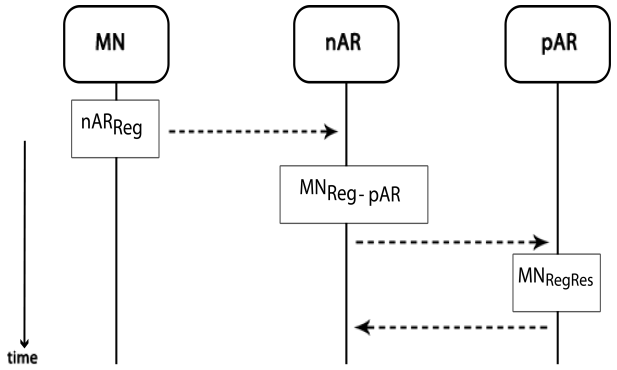


FIGURE 6. Ticket collection from a previous AR.

2) The nAR receives the message and sends a message ($MN_{Reg-pAR}$) to the pAR (previously attached AR of the MN), which has K_{MN-AR}.

$$nAR- > pAR : MN_{Reg-pAR}(E(K_{pAR-nAR}, ID_{MN}||N_{MN}+1))$$

3) The pAR sends the message (MN_{RegRes}) back to the nAR by adding the information about the time for which the ticket is valid and K_{TK}. The pAR encrypts this using the key it shares with the nAR.

$$pAR- > nAR : MN_{RegRes}(E(K_{pAR-nAR}, ID_{MN}||TT||K_{TK}))$$

The nAR decrypts the message to get the ticket using the key it shares with the pAR. Once, it gets K_{TK}, it decrypts TK and extracts K_{MN-AR}. The flow of the messages is shown in Figure 6.

In this section, we have proposed a protocol which makes the proposed location management secure. We have illustrated how an MN interacts with an AR and how a ticket is generated by the AS for a user session. We have also illustrated how a ticket can be reused to reduce authentication latency. In the next section, we evaluate the performance of TALM.

IV. PERFORMANCE EVALUATION

In this section, we compare TALM with TAP-TLS. The EAP-TLS based authentication performs EAP exchange between an MN and its AS; there is no concept of ticket reuse in EAP-TLS. Typically a successful TAP-TLS authentication has four steps of authentication process [35]. For the authentication related performance metrics, the simulation program is written in C++ and the simulation time is one year. We also calculate the average LQ processing time using the analytical model. The performance metrics are listed below:

1) Authentication latency (L_{auth}): Authentication latency is the latency to create a ticket (by an MN to authenticate itself to the AR for the first time), or renew a ticket when it is expired, or regain a ticket (from previous AR) in case of an IP handover. IP handover occurs when an MN changes its point of network access. The average authentication latency is calculated by dividing the cumulative authentication latency over a period of time by the total number of location update messages in that time period.
2) Number of authentication messages per location update message (N_{auth}): The number of authentication messages (requiring TALM protocol steps) in a time period is the number of times an MN authenticates itself to an AR in cases of ticket creation, ticket renewal, and ticket collection (from previous AR) in that time period. The number of authentication messages per location update message is calculated by dividing the total number of authentications (performed by an MN) by the total number of location update messages (sent by an MN) within a time period.
3) Average LQ processing time (Θ): It is the average response time of an LQ.

A. CALCULATION OF AUTHENTICATION LATENCY FOR TALM

Let *nhops* be the number of hops between the AR and the AS, T_{MN-AR} be the message transmission delay on one hop, T_{AR-AS} be the message transmission delay from the AR to the AS, and T_{MN-AS} be the message transmission delay from the MN to the AS.

We use the same parameter settings as defined in [35]. We select the value of *nhops* in the interval [5, 9], $T_{MN-AR} = 20\ ms$, the velocity of an MN is taken in the interval $[10, 40]m/s$, and the radius of the AR-area for an MN in the interval $[100, 200]m$.

There can be n hops between an AR and an AS. So, the message transmission delay from the AR to the AS is $T_{AR-AS} = nhops \times T_{MN-AR}$.

For MN to AS message transmission, the delay is divided into two parts. First, when a message is sent by an MN to the AR and second, when the message is sent to the AS from an AR. This transmission delay is $T_{MN-AS} = T_{MN-AR} + T_{AR-AS} = 20\ ms + (nhops \times T_{MN-AR})$.

Following is the order of messages for a ticket creation:

1) $MN- > AR$; 1 hop, $0.02s$
2) $AR- > AS$; n $hops$, $n \times 0.02s$
3) $AS- > AR$; n $hops$, $n \times 0.02s$
4) $AR- > MN$; 1 hop, $0.02s$
5) $MN- > AR$; 1 hop, $0.02s$

As the message transmission delay per hop is $0.02s$, the number of hops between an MN and the AR is 1, and the number of hops between an AR and the AS is *nhops* so, the time for the complete authentication process $= T(MN- > AR) + T(AR- > AS) + T(AS- > AR) + T(AR- > MN) + T(MN- > AR) = 0.02 + (nhops \times 0.02) + (nhops \times 0.02) + 0.02 + 0.02$s.

Following is the order of messages for a ticket creation:

1) $MN- > nAR$; 1 hop, $0.02s$
2) $nAR- > pAR$; 1 hop, $0.02s$
3) $pAR- > nAR$; 1 hop, $0.02s$

The ticket collection time (from the previous AR) = $T(MN->nAR)+T(nAR->pAR)+T(pAR->nAR)=0.02+0.02+0.02$s.

B. CALCULATION OF AUTHENTICATION LATENCY FOR TAP-TLS

The initial authentication delay is the authentication delay when an MN creates or renews (on its expiry) a ticket. As calculated in [35], it is $3T_{MN-AR}+T_{AR-AS}+T(m,T_{MN-AS})$, where T_{MN-AR} is the transmission delay between an *MN* and an *AR*, T_{AR-AS} is the transmission delay between an *AR* and the *AS* and $T(m,T_{MN-AS})$ is a function of T with $m=4$, and the transmission delay between an *MN* and the *AS* is $T_{MN-AR}+T_{AR-AS}$.

As calculated in [35], the handover authentication delay in EAP-TLS $=L_{l2}+$ initial authentication delay $+D_{SA}+D_{RS}+D_{REG}+D_{P_{LMA}}$, where $L_{l2}=0.04535$, $D_{SA}=4T_{MN-AR}$, D_{RS} is the transmission delay between an MN and an AR, $D_{REG}=2T_{AR-LMA}$, and $D_{P_{LMA}}=T_{AR-AS}+T_{MN-AR}$.

C. CALCULATION OF THE AVERAGE LQ PROCESSING TIME

Our choice of the parameter settings are comparable in magnitude with the ones defined in [11], [12], and [35]. We select λ_{LSQ} in the interval $\left[\frac{5}{3600},\frac{50}{3600}\right]$, $\lambda_{LUQ}=\frac{1}{3600}$, $N_{MN}=30\times10^6$, N_{AN} in the interval $\left[30,3\times10^7\right]$, $R_{delay}=20\ ms$, and N_{hops} in the interval [5, 10].

In order to compute the values of μ_i and μ_f, we implement an *LR* (that is maintained on an AN and is part of the DHT) as an in-memory database (IMDB) as well as the finger-table (also in the memory) of the AN for the DHT overlay routing. Main memory databases provides quicker access than the disk-optimized databases because a disk access is slower than a memory access. Accessing data in-memory eliminates the seek time when querying the data, which, consequently results in faster and more predictable performance than disk [37]. Also, the main memory sizes have grown big, so implementing an LR as an IMDB has become realistic.

The implementation is done on a standard hardware — Intel Core i3 CPU *M*390 with clock speed 2.67 *GHz* and 4 *GB* main memory — and run on Linux OS. By taking the average of 100, 000 measurements, $\mu_i=2\times10^{-6}$ and $\mu_f=11\times10^{-6}$. The dominant part of the average LQ processing (response) time is the standard routing delay in Internet. It is due to the in-memory processing of the LR and the finger-table of an AN. In the next section, we take λ_{LSQ}, N_{AN}, and N_{hops} as variable parameters and show the results for Θ besides showing the results for the average authentication latency and the average number of authentication messages per location update in TALM.

V. RESULTS AND DISCUSSION

We now present the network topology, the parameter settings, the results of the simulations, and our discussion of the results.

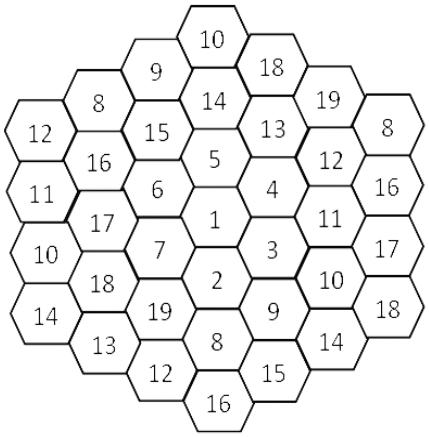


FIGURE 7. Cellular network topology.

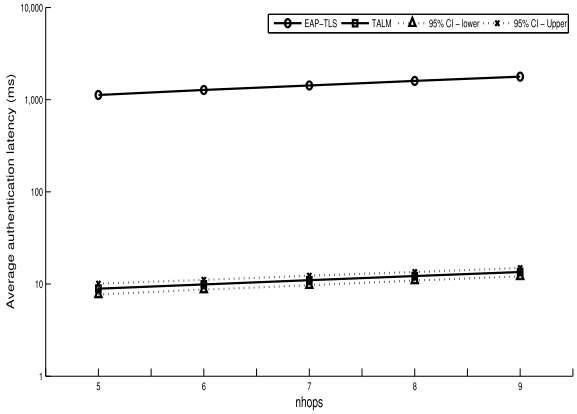


FIGURE 8. L_{auth} vs. *nhops*.

A wrapped-around cellular network topology of 19 hexagonal ANs is used in the simulations as shown in Figure 7.

For the simulations, we choose the AN-Area residence time α to be an exponentially distributed random variable as also chosen in [12] and [38]–[40]. The results of the seven scenarios shown in Table 3 are given and discussed in the following sections.

A. AVERAGE AUTHENTICATION LATENCY

We compare the average authentication latency of TALM with that of EAP-TLS. The variable parameter of this experiment is *nhops*. The results of this experiment are shown in Figure 8. With the increase in *nhops*, the average authentication latency for both the protocols increases. The results show that TALM performs better than TAP-TLS. We calculate the simulation results of TALM with 95% confidence interval as shown in Figure 8. The percentage reduction in authentication latency by TALM is nearly equal to 99%. This is mainly due to the reuse of the authentication ticket in TALM.

B. AVERAGE NUMBER OF AUTHENTICATION MESSAGES

Figure 9 shows the average number of authentication messages per location update as the lifetime of a ticket varies. With the increase in the lifetime of the ticket, the average number of authentication messages per location

TABLE 3. Scenarios and parameter settings.

Scenario	Metric	Variable parameter	Parameter values
1	L_{auth}	*nhops*	Ticket lifetime = 24 hrs, *nhops* $\in$ {5,6,7,8,9}, Velocity = 1 m/s, AN-Area = 10^4 km^2
2	N_{auth}	Ticket lifetime	Ticket lifetime $\in$ {12,24,48,72} hrs, *nhops* = random[5,9], Velocity = 1 m/s, AN-Area = 10^4 km^2
3	N_{auth}	Average velocity of MNs	Ticket lifetime = 24 hrs, *nhops* = random [5,9], Velocity $\in$ {1,5,10} m/s, AN-Area = 10^4 km^2
4	N_{auth}	AN-Area	Ticket lifetime = 24 hrs, *nhops* = random [5,9], Velocity = 1 m/s, AN-Area $\in$ {10^4,10^3,10^2,10,1} km^2

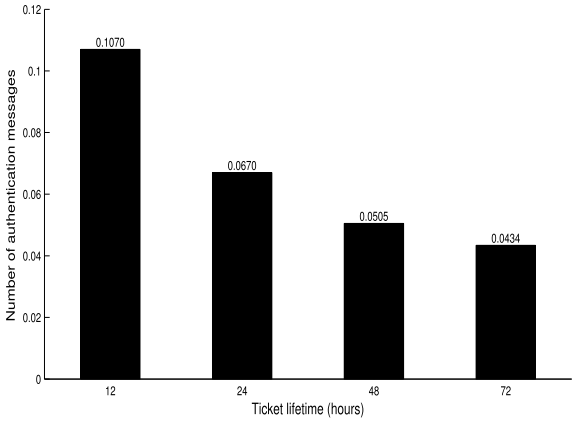


FIGURE 9. N_{auth} **vs. Ticket Lifetime.**

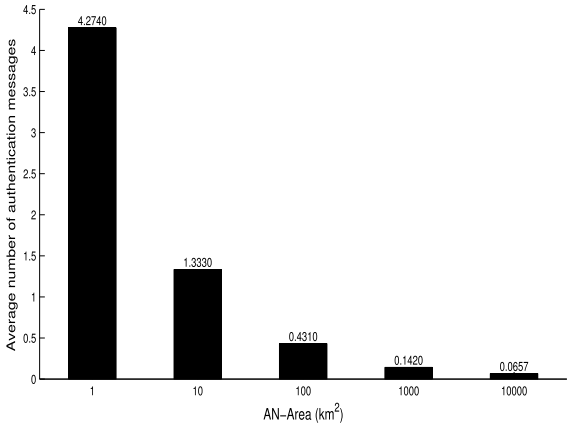


FIGURE 11. N_{auth} **vs. AN-Area.**

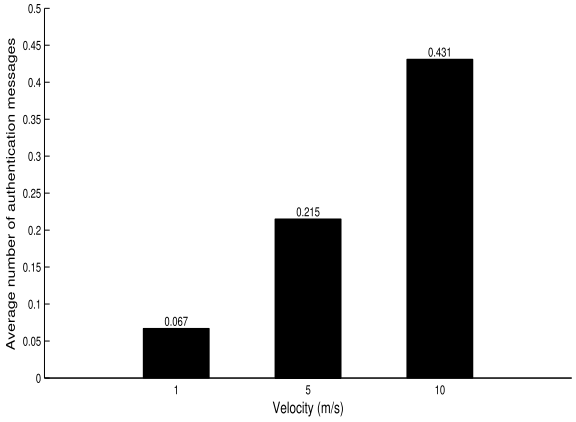


FIGURE 10. N_{auth} **vs. Velocity.**

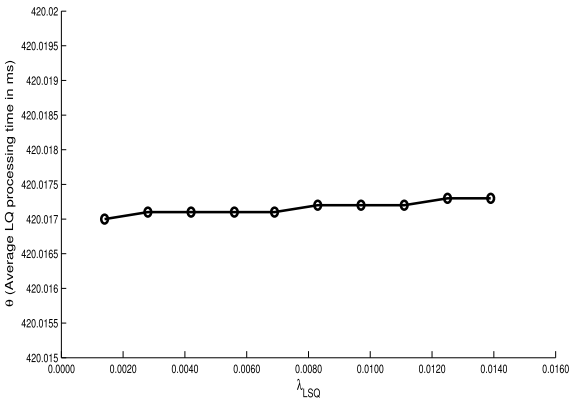


FIGURE 12. Θ **vs.** λ_{LSQ}.

update decreases. For bigger lifetime of a ticket, an MN faces less number of authentication ticket expiries than the number of the expiries when the lifetime is less.

Figure 10 shows the effect on average number of authentication messages per location update when the velocity of an MN is varied. We can see that the average number of authentication messages increases with the increase in the velocity. This is due to the reason that with the increase in velocity, the MN experiences more IP handovers. Hence, the MN performs more handover authentications and it results in the increase of the average number of authentication messages.

Figure 11 shows the effect on average number of authentication messages per location update when the area of an AN is varied. We observe that the average number of authentication messages decreases with the increase in the AN-Area. This is due to the reason that with the increase in the AN-Area, an MN experiences less IP handovers. Hence, the MN performs less number of handover authentications and it results in the decrease of the average number of authentication messages.

C. AVERAGE LQ PROCESSING TIME

We vary λ_{LSQ} in the interval $\left[\frac{5}{3600}, \frac{50}{3600}\right]$, choose $N_{hops} = 7$, $N_{AN} = 30$ and calculate Θ using Equation (6). The results are shown in Figure 12. The results show that there is only a little increase in Θ as we increase λ_{LSQ} but keep N_{hops} and N_{AN} constant. It is due to the scalability of the proposed distributed location management scheme using the Chord circle.

Then, we vary N_{hops} in the interval [5, 10], choose $\lambda_{LSQ} = 0.0014$, $N_{AN} = 30$ and calculate Θ using Equation (6).

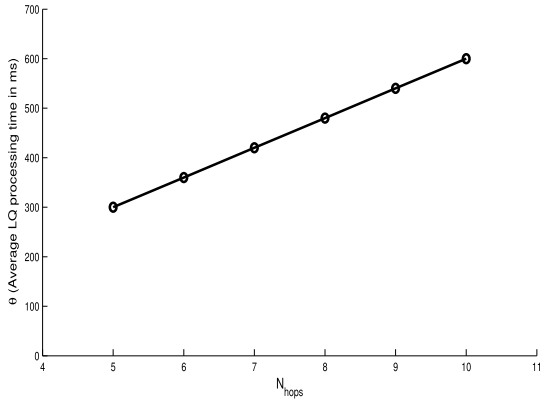


FIGURE 13. Θ vs. N_{hops}.

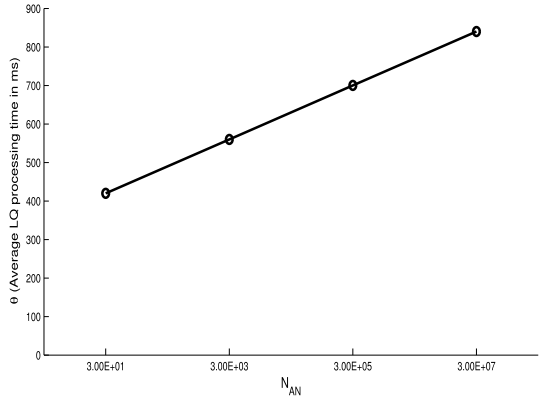


FIGURE 14. Θ vs. N_{AN}.

The results are shown in Figure 13. The results show that Θ is a linear function of N_{hops}, when λ_{LSQ} and N_{AN} are kept constant.

Last but not the least, we vary N_{AN} in the interval $[30, 3 \times 10^7]$, choose $\lambda_{LSQ} = 0.0014$, $N_{hops} = 7$ and calculate Θ using Equation (6). The results are shown in Figure 14. The results show that Θ is a linear function of N_{AN}, when λ_{LSQ} and N_{hops} are kept constant.

VI. CONCLUSION

In this paper, we have proposed a new mechanism for secure and fault-tolerant operations for DMM. The proposed mechanism utilizes a distributed hash table of access nodes for distributed location management. It is mapping the network prefixes of the access nodes and the location binding information of the MNs on a Chord circle. A replication strategy has been described for the proposed mechanism. The distribution of location server is important as it addresses the limitations and weaknesses of a central location server like single point of failure and attack, traffic bottleneck resulting in long delays in the additions and update of location binding information of MNs, and long response times of location queries that are sent by the CNs. The analytical model that has been used to calculate the average processing (response) time of a location query can help in designing and planning the processing requirements on an AN in future 5G networks.

The proposed mechanism also adopted a ticket-reuse approach called TALM for improving authentication performance while providing secure authentication of MNs. It provides security to the location additions and updates on the wireless link between an MN and its AR. The work is important as it reduces the latency of the authentication process and the number of authentication messages that are sent for a location update message. The results demonstrate significant performance gain of the proposed protocol as compared to a non-ticket based existing wireless authentication protocol, TAP-TLS. In order to replace all centralized components with their distributed counterparts for future 5G networks, it is desirable to have a distributed authentication service for secure handovers.

ACKNOWLEDGEMENTS

The authors are thankful to Prof. Renato Lo Cigno (*DISI*, University of Trento, Italy) for his guidance related to the analytical modeling part of the manuscript.

REFERENCES

[1] D. Johnson, C. Perkins, and J. Arkko, *Mobility Support for IPv6*, document IETF RFC 3775, 2004.
[2] S. Gundavelli, K. Leung, V. Devarapalli, K. Chowdhury, and B. Patil, *Proxy Mobile IPv6*, document IETF RFC 5213, 2008.
[3] *The IETF DMM WG*. Accessed: Dec. 22, 2017. [Online]. Available: http://datatracker.ietf.org/wg/dmm/
[4] T.-X. Do and Y. Kim, "Control and data plane separation architecture for supporting multicast listeners over distributed mobility management," *ICT Exp.*, vol. 3, no. 2, pp. 90–95, Jun. 2017.
[5] L. Guo, J. Wu, G. Li, J. Li, and J. Wu, "Software-defined networking model for smart transformers with ISO/IEC/IEEE 21451 sensors," *ICT Exp.*, vol. 3, no. 2, pp. 67–71, Jun. 2017.
[6] J. Kim, D. Kim, and S. Choi, "3GPP SA2 architecture and functions for 5G mobile communication system," *ICT Exp.*, vol. 3, no. 1, pp. 1–8, Mar. 2017.
[7] P. Zhang, J. Lu, Y. Wang, and Q. Wang, "Cooperative localization in 5G networks: A survey," *ICT Exp.*, vol. 3, no. 1, pp. 27–32, Mar. 2017.
[8] R. Sharan Sinha, Y. Wei, and S.-H. Hwang, "A survey on LPWA technology: LoRa and NB-IoT," *ICT Exp.*, vol. 3, no. 1, pp. 14–21, Mar. 2017.
[9] H.-A. Chan, H. Yokota, J. Xie, P. Seite, and D. Liu, "Distributed and dynamic mobility management in mobile Internet: Current approaches and issues," *J. Commun.*, vol. 6, no. 1, pp. 4–15, 2011.
[10] J.-H. Lee, J.-M. Bonnin, P. Seite, and C. H. Anthony, "Distributed IP mobility management from the perspective of the IETF: Motivations, requirements, approaches, comparison, and challenges," *IEEE Wireless Commun.*, vol. 20, no. 5, pp. 159–168, Oct. 2013.
[11] H. Ali-Ahmad, K. Munir, P. Bertin, K. Guillouard, M. Ouzzif, and X. Lagrange, "Processing loads analysis of distributed mobility management and SIP-based reachability," *Telecommun. Syst.*, vol. 63, no. 4, pp. 681–696, 2016.
[12] K. Munir, X. Lagrange, P. Bertin, K. Guillouard, and M. Ouzzif, "Performance analysis of mobility management architectures in cellular networks," *Telecommun. Syst.*, vol. 59, no. 2, pp. 211–227, 2015.
[13] P. Bertin, S. Bonjour, and J.-M. Bonnin, "A distributed dynamic mobility management scheme designed for flat IP architectures," in *Proc. 3rd Int. Conf. New Technol., Mobility Secur.*, Nov. 2008, pp. 1–5.
[14] J. Xie and I. F. Akyildiz, "A novel distributed dynamic location management scheme for minimizing signaling costs in mobile IP," *IEEE Trans. Mobile Comput.*, vol. 1, no. 3, pp. 163–175, Mar. 2002.
[15] R. Prakash, Z. Haas, and M. Singhal, "Load-balanced location management for cellular mobile systems using quorums and dynamic hashing," *Wireless Netw.*, vol. 7, no. 5, pp. 497–512, 2001.
[16] G. Krishnamurthi, M. Azizoglu, and A. K. Somani, "Optimal distributed location management in mobile networks," *Mobile Netw. Appl.*, vol. 6, no. 2, pp. 117–124, 2001.

[17] H. Nguyen-Minh and H. R. van As, "A comparative study on distributed location management strategies in wireless networks," *Pers. Wireless Commun.*, vol. 51, pp. 111–122, Sep. 2000.
[18] K. Lee, H. W. Lee, S. Jha, and N. Bulusu, "Adaptive, distributed location management in mobile, wireless networks," in *Proc. IEEE Int. Conf. Commun.*, vol. 7. Jun. 2004, pp. 4077–4081.
[19] M. Song, R.-J. Feng, J.-W. Huang, and J.-D. Song, "A distributed location management strategy for next-generation IP-based wireless networks," *J. China Univ. Posts Telecommun.*, vol. 13, no. 3, pp. 38–42, 2006.
[20] Y. Xiao, Y. Pan, and J. Li, "Design and analysis of location management for 3G cellular networks," *IEEE Trans. Parallel Distrib. Syst.*, vol. 15, no. 4, pp. 339–349, Apr. 2004.
[21] B. Sidhu and H. Singh, "Location management in cellular networks," *World Acad. Sci., Eng. Technol.*, vol. 25, pp. 314–319, Jan. 2007.
[22] K. Kyamakya and K. Jobmann, "Location management in cellular networks: classification of the most important paradigms, realistic simulation framework, and relative performance analysis," *IEEE Trans. Veh. Technol.*, vol. 54, no. 2, pp. 687–708, Mar. 2005.
[23] C. Selvan, R. Shanmugalakshmi, and V. Nirmala, "Location management technique to reduce complexity in cellular networks," *Int. J. Comput. Sci. Issues*, vol. 7, p. 1, Jul. 2010.
[24] Y. J. Zhai, X. Y. Mao, Y. Wang, J. Yuan, and Y. Ren, "A DHT-based fast handover management scheme for mobile identifier/locator separation networks," *Sci. China Inf. Sci.*, vol. 56, no. 12, pp. 1–15, 2013.
[25] M. Bezahaf, L. Iannone, M. D. de Amorim, and S. Fdida, "Transparent and distributed localization of mobile users in wireless mesh networks," in *Proc. Int. Conf. Heterogeneous Netw. Quality, Rel., Security Robustness*, Berlin, Germany, 2009, pp. 513–529.
[26] I. M. R. Stoica, D. Karger, M. F. Kaashoek, and H. Balakrishnan, "Chord: A scalable peer-to-peer lookup service for Internet applications," in *Proc. Conf. Appl., Technol., Archit., Protocols Comput. Commun. (SIGCOMM)*, New York, NY, USA, 2001, pp. 149–160.
[27] W. A. Imtiaz, M. Afaq, and M. A. K. Babar, "mSCTP based decentralized mobility framework," *Int. J. Adv. Comput. Sci. Appl.*, vol. 2, no. 9, pp. 106–112, 2011.
[28] W. A. Imtiaz, "Two-tier CHORD for decentralized location management," *Int. J. Comput. Appl.*, vol. 69, p. 4, Jan. 2013.
[29] H. Ali-Ahmad, K. Munir, X. Lagrange, M. Ouzzif, and P. Bertin, "Analysis of SIP-based location updates in distributed mobility management schemes," in *Proc. 17th Int. Conf. Innov. Mobile Internet Services Ubiquitous Comput.*, 2013, pp. 13–18.
[30] D. Simon, B. Aboba, and R. Hurst, *The EAP-TLS Authentication Protocol*, document RFC 5216, Mar. 2008.
[31] H. Zhou, H. Zhang, and Y. Qin, "An authentication method for proxy mobile IPv6 and performance analysis," *Secur. Commun. Netw.*, vol. 2, no. 5, pp. 445–454, 2009.
[32] M. Naor and U. Wieder, "Know thy neighbor's neighbor: Better routing for skip-graphs and small worlds," in *Proc. Peer-to-Peer Syst.*, 2005, pp. 269–277.
[33] G. S. Manku, "The power of lookahead in small-world routing networks," in *Proc. STOC*, 2004, p. 1.
[34] G. S. Manku, M. Naor, and U. Wieder, "Know thy neighbor's neighbor: The power of lookahead in randomized p2p networks," in *Proc. 36th Annu. ACM Symp. Theory Comput.*, 2004, pp. 54–63.
[35] J.-H. Lee and J.-M. Bonnin, "HOTA: Handover optimized ticket-based authentication in network-based mobility management," *Inf. Sci.*, vol. 230, pp. 64–77, May 2013.
[36] R. Steinmetz and K. Wehrle, *P2P Systems and Applications* (Lecture Notes in Computer Science). Berlin, Germany: Springer, 2005, pp. 95–117.
[37] Q. Kang, C. Jin, Z. Zhang, and A. Zhou, *MemTest: A Novel Benchmark for In-memory Database*(Lecture Notes in Computer Science). Cham, Switzerland: Springer, 2014.
[38] R. G. Akl, M. V. Hegde, and M. Naraghi-Pour, "Mobility-based CAC algorithm for arbitrary call-arrival rates in CDMA cellular systems," *IEEE Trans. Veh. Technol.*, vol. 54, no. 2, pp. 639–651, Feb. 2005.
[39] D. Hong and S. S. Rappaport, "Traffic model and performance analysis for cellular mobile radio telephone systems with prioritized and nonprioritized handoff procedures," *IEEE Trans. Veh. Technol.*, vol. VT-35, no. 3, pp. 77–92, Aug. 1986.
[40] R. Thomas, H. Gilbert, and G. Maziotto, "Influence of the moving of the mobile stations on the performance of a radio mobile cellular network," in *Proc. 3rd Nordic Seminar Digit. Land Mobile Radio Commun.*, 1998, pp. 12–15.

KASHIF MUNIR received the Ph.D. degree from the University of Innsbruck, Austria, in 2009. He was a Post-Doctoral Researcher with IMT Atlantique (formerly known as TELECOM Bretagne), France, from 2011 to 2012. He is currently an Associate Professor with the Department of Computer Science, National University of Computer and Emerging Sciences, Islamabad, Pakistan. He has authored of numerous peer-reviewed conference and journal publications. His areas of research interests include admission and congestion control, the quality of service for bulk data transfers, the performance modeling of computer and communication systems, high-performance computing, and mobility cost analysis.

EHTESHAM ZAHOOR received the Ph.D. degree from the Université de Lorraine, France, for the research he carried out at the Lorraine Research Laboratory in Computer Science and its Applications (LORIA/INRIA). He was an ATER with the Université de Lorraine. He has been with the National University of Computer and Emerging Sciences, Islamabad, Pakistan, as an Assistant Professor, since 2012. He has been an active Researcher in the domain of large-scale distributed systems and an author of numerous peer-reviewed conference and journal publications. He also has vast experience in the software industry.

RAFIA RAHIM received the M.S. degree in computer science from the National University of Computer and Emerging Sciences (NUCES), Islamabad, Pakistan, in 2017. She is currently a Lecturer with the Department of Computer Science, NUCES. Her research interests include mobile networks, information security, the performance evaluation of computer and communications systems, data science, and machine learning.

XAVIER LAGRANGE (SM'08) received the Degree in engineering from the Ecole Centrale des Arts et Manufactures, Paris, France, in 1984, and the Ph.D. degree from Télécom ParisTech in 1998. He is currently a Professor with IMT Atlantique, France, and the Head of the research group ADOPNET, IRISA. He has co-authored of several text books in French on wireless networks. His research interests include resource allocation, medium access control, and performance analysis for 4th and 5th generation cellular networks.

JONG-HYOUK LEE received the Ph.D. degree in computer engineering from Sungkyunkwan University, Suwon, South Korea. He was with INRIA, France, for IPv6 vehicular communication and security research. He was an Assistant Professor with TELECOM Bretagne, France. In 2013, he moved to Sangmyung University, Cheonan, South Korea. He is an author of the Internet Standards: IETF RFC 8127 and IETF RFC 8191. His research interests include blockchain, malware analysis, and protocol analysis. He received the Best Paper Award from the IEEE WiMob 2012, the 2015 Best Land Transportation Paper Award from the IEEE Vehicular Technology Society, and the Haedong Young Scholar Award in 2017. He was a tutorial speaker at the IEEE WCNC 2013, the IEEE VTC 2014 Spring, and the IEEE ICC 2016. He was introduced as the Young Researcher of the Month by the National Research Foundation of Korea Webzine in 2014.