A Novel Method in Mobile Subscriber Location Estimation for Enhancement of Paging Procedure in Mobile Cellular Networks

Kieu-Ha Phung*, Van-Phu Tran, Minh-Hung Le, Nguyen Tai Hung, Nguyen Huu Thanh
School of Electrical and Electronic Engineering, Hanoi University of Science and Technology
Ha Noi, VietNam
*ha.phungthikieu@hust.edu.vn

Pham Quang Diep, Le Van Hien, Mai Xuan Ngoc, Le Minh Tu
Viettel Hi-Technology Industries Corporation,
Hanoi, Vietnam

***Abstract*—Mobility management is a critical aspect of cellular network operations, and paging is an essential procedure of this process. The emergence of 5G and beyond networks serving huge number of mobile users/devices requiring strict QoS-required, heterogeneous, has increased the complexity of mobility management. An highly efficient paging procedure demands the fast response of mobile users' location information in the minimal expense of network resources. The investigation of data collected from the mobile network operation has posed that the variation of wireless communication has high impact on the efficiency of the paging procedure. An improvement to paging procedure has been proposed and evaluated with the data collected from a real-world operated mobile network. The results have shown that the number of radio base station joining in the paging procedure can be reduced up to about 95% with the conservation of the required paging success ratio, over 95%.**

***Index Terms*—Cellular Networks, Mobility Management, Paging procedure**

I. INTRODUCTION

Mobility management is one of the essential network functions of 4G/5G mobile cellular networks. It is to ensure that mobile devices can move among coverage areas without interruption of call connection and data transmission. The primary tasks of mobility management include locating mobile user equipment (UE) via Tracking Area Update (TAU) or Paging procedures [1]. The Mobility Management Entity in 4G network (4G-MME) or the Access and Mobility Management Function in 5G network (5G-AMF) will handle the tasks. 4G-MME or 5G-AMF has the information of the Tracking Area (TA) or the Location Area (LA) where the UEs are registered. However, this information might not reflect UEs' current locations when the UEs are in idle state. To locate the idle users when incoming services are requesting (e.g. incoming voice call, data call, etc...), the network initiates a Paging procedure [1]. 4G-MME or 5G-AMF estimates and decides the list of *paging e/g-NBs* which are the radio base station sending paging messages. The precise detection of the location of user in an idle state is critical to ensure the service continuation and service quality. For brevity, we use MME to refer 4G-MME and 5G-AMF and eNB for evolved Node Base Station and 5G Base Station.

When an UE recognized to be in idle state has service requests targeted at, the paging procedure is triggered at MME. A *page message* containing the UE identifier is broadcasted by eNBs. The broadcast request is assigned to a cluster of radio cells that together form a location area (LA) where the UE is expected to be inside. The last reported position of the UE, inferring from the registered eNB, normally identifies the location area to perform the message broadcast procedure. The UE monitors the page messages transmitted by the eNB of the radio cell in which it is located and, on detecting its own identifier, responds by transmitting a page response message to the eNB. Communication is then established between the MME and the UE via the eNB that received the page response message.

The number of eNBs receiving page message from MME is a critical parameter. A large number of page eNBs can help to quickly reach the UEs, however, will induce the MME signalling overhead in the network system.

An idle-state UE mobility management basically involves a 3-stage paging procedure [1]. The first stage is triggered upon receiving a connection request destined an idle state user. The MME transmits paging message only to the eNB where the UE was connected last. If the UE is not found there then in the second stage, the MME broadcasts the paging message to all eNBs within the TA of such previously connected eNB and waits for a response. If this also fails then at the final stage, the paging message is broadcasting across all the eNBs in the TA List (TAL) of the UE, assigned by the MME.

The mobility management tasks accounts to nearly one third of MME signaling overhead in the real network deployments [2]. The evolution towards small cells and pico cells in 5G and beyond networks will further aggregate the overhead. In the meanwhile, several QoS requirements, page response time and connection establishment delay, should be carefully considered in order to support new 5G and beyond use cases (e.g., edge computing, high speed mobility, and real time gaming) and

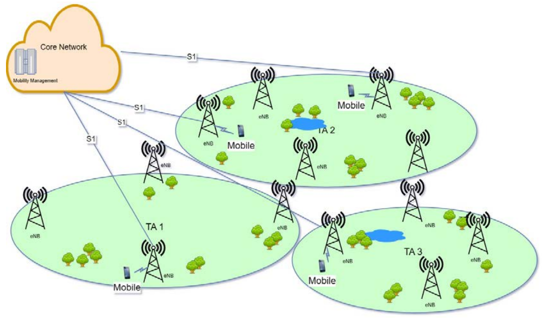

Fig. 1: Tracking Area in cellular mobile network

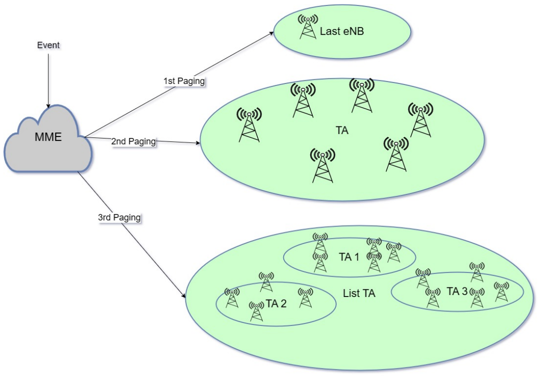

Fig. 2: 3-stage Paging Procedure

increasing number of mobile users.

II. Related Work in Paging Procedure Enhancement

In a 4G/ 5G network, the paging procedure is performed through activation notification transmission to UEs. It is accomplished by dividing time into paging frames, sending paging notifications within paging occasions, and utilizing a paging record to determine which mobile devices need to listen for the notification [3]. When a mobile device receives a paging notification, it responds and activates itself to receive services within the 5G network. There are several approaches which has been recommended in [1],[3] for a 5G network to reduce the signalling overhead in increasing demands of mobile connectivity.

- Paging Grouping: Group mobile devices with similar activation requirements which can be paged in the same paging occasion to reduce paging traffic and save energy.
- Flexible Paging Configuration: Adjust the paging process based on network changes and resource utilization requirements to improve performance and resource efficiency.
- Prioritized Paging: Assign priority levels to paging processes. Important or time-sensitive requests can be prioritized during paging, ensuring quick handling of critical requests.
- Enhanced Paging Indication: Provide additional indications in the paging notification to provide more detailed information to the mobile devices. This can help minimize the time and resources required for device identification and activation.
- Paging Area Optimization: Optimize the paging area by adjusting its size and shape to meet specific network requirements. This can minimize unnecessary paging traffic and optimize device activation.

Several approaches have been suggested to minimize signaling load in TAU (Tracking Area Update) and Paging strategies by taking into account time, movement, and distance based on the position of user devices (UE) [4]. One of the proposed methods is the overall Paging approach (Blanket Paging), where all Tracking Areas (TA) are simultaneously paged to determine the UE's location. However, this method is ineffective and demands high bandwidth [5]. Another method proposed in [6] is sequential paging, where UE information is maintained, and Paging is performed for a group of cells based on the probability of finding the UE within them. Nevertheless, this method results in significant Paging latency, hence, increasing connection establishment time, particularly under heavy load conditions. To overcome these limitations, a region-based Paging solution with variable sizes has been proposed [7]. This solution defines three different dimensions for Paging regions: small, medium, and large. Depending on the time elapsed since the last UE service, an appropriately sized Paging region is paged simultaneously. In summary, these methods aim to optimize the Paging process and reduce signaling load by considering various factors like UE position, elapsed time, and the likelihood of finding the UE in specific areas.

Balancing the TAU and Paging procedure is an optimization problem aimed at achieving a better Tracking Area List (TAL). A multi-objective optimization algorithm based on Population Decomposition strategy allows the creation of an improved TAL [8]. This algorithm integrates geographic information that impacts user mobility to generate an optimal TAL plan. However, the TAL optimization model is only applicable during network deployment. Another advanced idea is to utilize a dynamic mapping of cells to TAs and TAs to TAL based on user movement [9]. This concept assists in reorganizing the TAL map based on user mobility within the network. Research [10] incorporates the concept of MME pooling into the optimization problem and compares the performance between a centralized MME system (one MME controlling one TAL) and a distributed MME system (one MME controlling one TA). The research findings indicate that the centralized MME system performs better, as the distributed system incurs additional costs due to TA relocations. Although the idea of distributed MME aligns with the flat architecture for next-generation mobile networks [11], it faces challenges in determining user locations from distributed databases.

In [12], probabilistic solutions utilizing methods based on Information Theory are proposed. In this approach, TAU techniques based on Bayesian and Entropy encoding are applied to minimize paging load. The idea is to use probabilities to estimate which UE is located in which area, and based on this information, paging decisions are made. The models have shown a significant reduction in signaling load, ranging from 60% to 80%. However, they also introduce additional computational load on each UE. This may not be practical in real-world deployments as the complex computational requirements on UEs can lead to high energy consumption or impact UE performance. In [13], a mobile management architecture with prediction capability is proposed. In this architecture, a data collection service application running on UEs periodically collects location and context information of UEs and shares this information with eNBs. This information is used as feedback to predict the UE's location and reduce signaling load in different user mobility scenarios. With the contribution of the service application and the sharing of location information, the system has the ability to predict potential UE location. As the UE moves, the system utilizes the collected and predicted information to minimize unnecessary Paging, thereby reducing signaling load in various user mobility cases.

With the explosion of deep learning with the ability to learn rules, to deal with unstructured data, and to synthesize information from learned data patterns, there have been many studies using these techniques with very good accuracy. However, deep learning techniques in general have too large computation and cannot be met in real time operation [14].

The issue of minimizing the number of paging eNBs while maintaining required call service QoS for UEs is interesting and critical. Our approach is to thoroughly analysis the data log of UEs when communicating with eNBs in each operation mode. The data can show the operation behaviour of an UE and with the aim of a low computation scheme that can be integrated to the existing paging procedure without resulting in high

III. Analysis of Mobile user data log

During the operation, UE and the system must follow a chain of procedures in order to function properly. The information related to each procedure is kept in the core systems, specifically at the Mobile Management Entity (MME) providing mobility session management. MME supports and managing the tracking area of the UE. Table 1 shows a sample of data log of an UE.

The explanation of information collected at an MME is follows.

- Procedure: show the specific procedure that happens.
- Result: shows the outcome state (successful or failed) after the procedure execution.
- Duration: show how long a specific UE was located in a certain cell.
- Retry: shows the number of attempts that the procedure is executed until the result is reported (successful or failed).

Procedure	Result	Duration	Retry	UE	TAC	Cell ID	Time
l-tau	success	791	0	**********73	12105	53888513	1665709214985
l-service-request	success	117	0	**********75	12105	53771524	1665709215002
l-handover	success	0	0	**********84	12105	61541122	1665709215005

Table 1 Several samples of data log collected at MME while UEs are connected to the network

- UE: is the unique mobile user equipment identifier.
- TAC (Tracking Area code): shows the unique identifier of the current tracking area where UE locates inside,
- Cell ID (Cell Identity): shows the identifier of the coverage area (cell) which is currently in communication with the UE. An eNodeB might control several cells and be responsible for all communication sessions with UEs of its cells. It is called the registered eNodeB.
- Timestamp: represent the time when the procedure happens. The resolution is minute based.

Figure 3 shows the three modes of UE, CONNECTED, IDLE and DE-REGISTERED, during the call services provided by a 4G/5G network. After the release of a session between UE and eNB, UE is in IDLE mode to save the energy cost. UE is resumed the CONNECTED mode after it submits a call request, or it calls on Tracking Area Update (TAU) successful during its movement. The paging procedure occurs when a connection request destined to such IDLE UE is submitted. The MME will issue the paging procedure to that destined UE. The eNBs that maintained the connection/communication in the previous sessions are called as the registered eNBs.

The three-stage paging procedure is shown in Fig. 2. The paging eNBs should base on the history of the registered eNBs, e.g. in the first stage, the paging eNB is the most recent registered eNB. The second stage, paging the set of eNBs surrounding the currently registered eNB to increase the successful paging probability, is called on when the first attempt is failed. Note the cluster of an eNB as the list of eNBs whose locations are inside the radius R from such eNB location. The larger cluster size, including the more number of eNBs, hence the higher paging traffic in the system. That is the trade-off in paging procedure as mentioned in advance.

Analysing the data log collected from the operation of UEs, (e.g. procedure event timestamp, connection status and registered eNBs), we observe several situations as follows. When UE is moving and using call services, the history of registered eNBs alters frequently, handover and TAU occurs, UE is connected, paging procedure is not called on. However, the variable nature of wireless communication because of

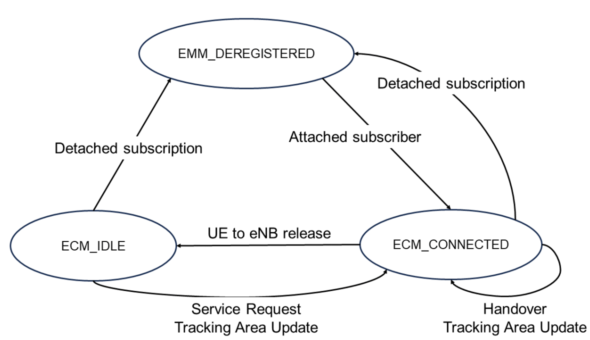

Fig. 3: State machine of UE connection to the mobile core system

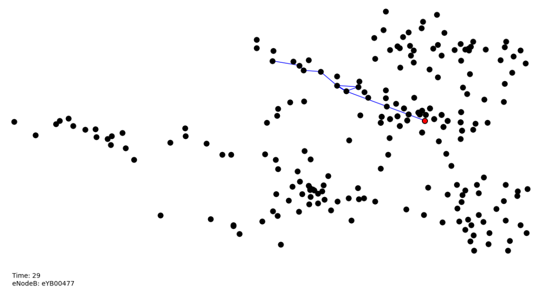

Fig. 4: Stable wireless communication

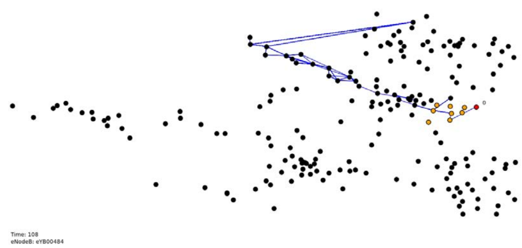

Fig. 5: Unstable wireless communication resulting in a registered eNB too far to be reachable in reality

multi-path fading, weather conditions, deployment environment, ect..., might cause the sudden changes of registered eNBs, i.e. the registered eNB might be too far to be reachable in a short period of time. It appears that the location of registered eNBs might not reflect the "real" location of UE. Hence, the error in paging procedure would be enormous when the paging decision is based on the registered eNBs appearing in such unstable communication state.

Figure 4 and Figure 5 show the track log of registered eNBs of an UE. The data is obtained from the real information collected from the mobile service operation of the cellular mobile network operated by Viettel company. In these figures, the black dots are the position of eNB station, the blue dots are the history eNBs and the red dot is the current register eNB. In Figure 4, the track log of an UE show its "feasible" travelling route through several eNBs. The figure 5 shows the sudden change issue, the position of a registered eNB is unreachable from the previously registered eNB in such short time interval between the two corresponding procedure events reported in the MME.

IV. Mobile User Location Estimation for Paging procedure improvement

We have defined UE's states related to the location estimation issue as follows:

- Stable communication: UE often connects to a single eNodeB in a relatively large period of time, showing that the UE might be not moving, or the UE has changed the registered/connected eNBs and the distance between the consecutive eNBs is rather large, showing that the UE might be moving.
- Unstable communication: UE changes rapidly the register/connected eNBs during a very short period of time but the physical distance between these consecutively registered eNBs is rather large. Conventionally, the distance is unfeasible to reach during such short period.

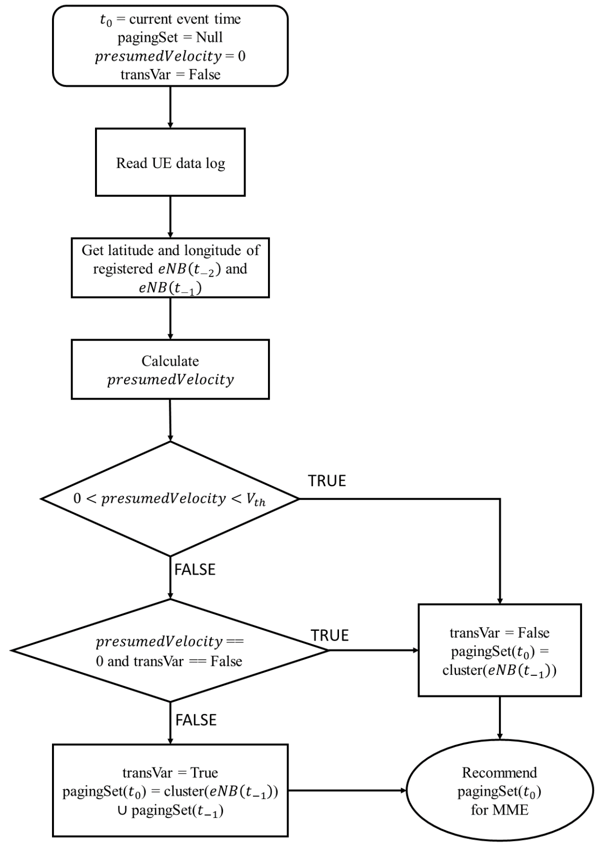

Fig. 6: The flow chart of the proposed scheme

We propose the procedure to detect the unstable communication state as shown in Fig.6. The timestamp of procedure events and the corresponding registered eNBs are used to assess the "presumed" velocity of the UE. The too large "presumed" velocity will imply an unstable communication. Hence, the paging eNB list decision should integrate this situation. The following section will present the scheme to detect the unstable communication and the corresponding paging eNB set to improve the succeed of paging procedure.

The summary of the variables, parameters and their definitions used in the algorithm is follows:

- t_0: current event,
- t_{-k}: event in the past, k steps (events) away from the current event,
- $timestamp(.)$: timestamp of the event,
- $dist(km)$: the distance between the two registered eNBs at $timestamp(t_0)$ and $timestamp_(t_{-1})$,
- V_{pre}: the presumed velocity to estimate of the UE's

displacement (the distance between the two eNBs) and the time interval, according to the equation [ref],

- V_{th}: threshold velocity (used to distinguish between UE moving or unstable transmission),
- *transVar*: state variable,
- *cluster(.)*: set of eNBs residing inside the ring of the radius of R(km) from an eNB,
- *pagingSet(.)*: the result of the algorithm, which is the set of eNBs that should be paged,
- R: the radius of the cluster (in kilometers).

$$V_{pre} = \frac{dist(eNB_1, eNB_2)}{Timestamp_2 - Timestamp_1} \quad (1)$$

V. Performance Evaluation and Discussion

The algorithm is implemented in Python. The simulation program will be executed with the input data collected from the service operation of mobile users in the Viettel mobile network in Vietnam. The raw data is normalized by the elimination of the UEs which has very few movement during the observed time. More specifically, the normalized data set contains the UEs which travel through at least 10 eNBs in the observed duration. The characteristic of data set and the set up parameters are shown in Table 2.

The assessment of the proposed algorithm is based on the "precision" of paging eNB list every time the paging procedure is called on. The decision is considered as precise when there is the "true" registered eNB in the obtained pagingSet. Based on the full information of the collected data set, the ground truth (the registered eNB of the paging-required UE at each paging procedure event) is to used to assess the outcomes of the algorithm. The evaluated parameter also includes the average number of paging eNBs. The assessment has been performed with the normalized data set.

The paging assessment based on the operation of the 4G network currently operated by Viettel mobile service provider in Vietnam is shown in Table 3. This assessment information

The number of UE	100
The observed time duration	7.00AM - 10.00AM
The number of eNBs in TAC	207
Cluster size (km)	3, 6, 9, 12, 15
Threshold velocity (km/second)	0.05, 0.1, 0.15, 0.2, 0.25

Table 2 The data set used for algorithm assessment and the investigation scenarios of cluster size and threshold velocity

Report time	First time PSR(%)	Overall PSR(%)
2023-04-20-00	96.62	98.33
2023-04-21-00	96.89	98.43
2023-04-22-00	96.92	98.44
2023-04-23-00	96.93	98.45
2023-04-24-00	96.92	98.44
2023-04-25-00	96.76	98.32
2023-04-26-00	96.8	98.36

Table 3 The Paging Success Ratio (PSR) in 4G mobile network operated by Viettel mobile service provider in Vietnam. The data is collected in the Tracking Area where the data set of UEs used to assess the proposed algorithm is collected.

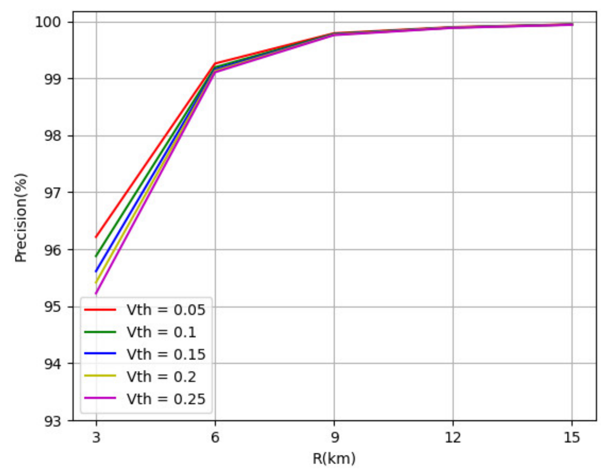

Fig. 7: The precision of the prediction with different cluster size (km) and threshold level of presumed velocity (km/s).

can be the reference to make the comparison with the result improvement obtained by the proposed algorithm since the Tracking Areas are identical. The results show that the overall Paging Succeed Ratio (PSR) of their paging procedure is reported at approximately 98% when the paging messages are transmitted to all eNBs in the observed Tracking Area. The paging messages are re-transmitted at the maximum of 3 times. The PSR of the first attempt is reported approximately 96.7%. It should be emphasized that the number of eNBs in the Tracking Area is rather large, about 200 eNBs. In Viettel mobile network deployment, this number can be one to several hundreds of eNBs per Tracking Area.

Figure 7 shows that the precision of the estimation is more than 99% when the cluster size is over 6km. When the cluster size is 3km, the precision can be in range of 95%-96%. The lower threshold of the presumed velocity results in the better precision. It means that the detection of "unstable communication" and the adjusted policy of paging eNB selection due to that unstable state has generated the better outcome. Figure 8 shows the expected outcome when the mechanism is deployed in the real-world system. The adjusted precision is approximately determined as the pagingSet precision times the first-time PSR of 96.8% in the real deployment, see Table 3.

Table 4 presents the adjusted precision of the estimation and the average number of paging eNBs when the threshold velocity is set at 0.1 km/s. It can be seen that at the cluster size of 10km, the scheme can help to achieve a paging success ratio at 93.02%. This evaluated results have already considered the failed communication. In terms of network traffic saving, the number of paging eNBs can be reduced to approximately 95%. An inconvenience of the proposed scheme is the usage of location information (latitude and longtitude) of eNBs for the state recognition and cluster's eNB member calculation.

There are two possibilities in using the proposed algorithm to enhance the efficiency of paging procedures.

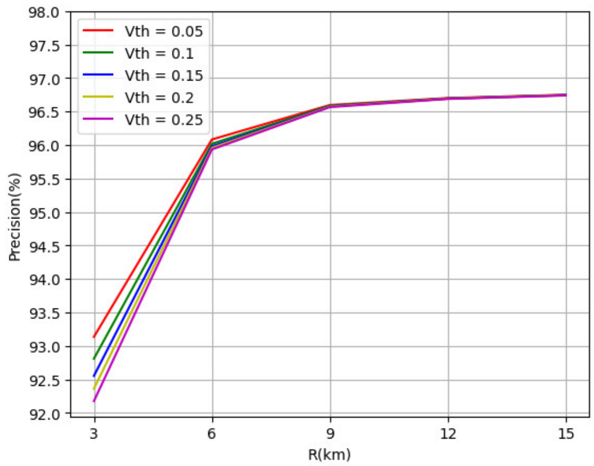

Fig. 8: The expected precision when the proposed method would be deployed in the system.

Cluster size	Average eNBs/paging	Precision (%)
3km	3.18	98.66
6km	7.96	99.7
9km	14	99.9
12km	20.74	99.95
15km	27.73	99.97

Table 4 $V_{th} = 0.1(km/s)$

- Online paging procedure: integrate in the system paging procedure and give online suggestion for the first attempt of paging message transmission.
- Offline recommendation: statistically analyse the data log of various UEs to study the optimal eNB cluster size for each Tracking Area. The differences in Tracking Area characteristic in terms on geographical region (e.g. rural area, suburban or urban area), eNBs distribution, UEs density, ect... would result in different optimal eNB cluster sizes. Such optimal cluster size would be used as a reference for listing the first-attempt paging eNBs without calling on the geographical locations of eNBs.

VI. Conclusion and Future Work

In mobile network, mobility management is essential in ensuring the service quality, especially when the service applications are demanding and the users are booming. Precise and fast response paging procedure is required, while lowering the traffic overhead created by the procedure. Our proposed algorithm for UE location estimate has shown that it can achieve the required paging succeed ratio with much lower number of paging eNBs, hence, the paging traffic can decrease drastically.

The future work will include the assessment of the proposed algorithm intensively with the data sets collected from various types of Tracking Areas, e.g. different in geographical regions, UEs' density, eNBs distribution. Further investigation in the system integration should be done to ensure practical efficiency requirements for real deployment.

References

[1] 3rd Generation Partnership Project (3GPP), "General Packet Radio Service (GPRS); Service description; Stage 2", number = "23.060", 2019.

[2] Nadine Akkari , Nikos Dimitriou, "Mobility management solutions for 5G networks: Architecture and services", vol.169, March 2020.

[3] Erik Dahlman, Stefan Parkvall, Johan Sköld, "5G NR: The Next Generation Wireless Access Technology", 2018.

[4] T. Deng, X. Wang, P. Fan, K. Li , "Modeling and performance analysis of a tracking-area-list-based location management scheme in LTE networks", IEEE Transactions on Vehicular Technology, vol.65, pp. 6417 - 6431, 2016.

[5] S. R. Parija, N. P. Nath, P. K. Sahu, S. S. Singh, "Dynamic profile based paging in mobile communication," International Conference on Microwave", Optical and Communication Engineering (ICMOCE), pp. 342-345, Oct 2015.

[6] R. H. Liou, Lin, Y. B. Lin, S. C. Tsai, "An Investigation on LTE Mobility Management", IEEE Transactions on Mobile Computing, vol.12, pp. 166-176, 2013.

[7] T. Hashimoto, T.Kubo, Y. Kitatsuji, " Lightweight paging based on location update intervals in LTE network", IEEE International Conference on Communications (ICC), pp. 1-6, May 2016.

[8] L. Chen, H. L. Liu, Z. Fan, S. Xie E. D. Goodman, "Modeling the tracking area planning problem using an evolutionary multi-objective algorithm", IEEE Computational Intelligence Magazine, vol.12 ,pp. 29-41, 2017.

[9] J. Moysen, M. Garcia-Lozano, "Learning-based Tracking Area List Management in 4G and 5G Networks", IEEE Transactions on Mobile Computing, 2019.

[10] E. Aqeeli, A. Moubayed, A. Shami, "Towards intelligent LTE mobility management through MME pooling", IEEE Global Communications Conference (GLOBECOM), pp. 1-6, Dec. 2015.

[11] S. Jeon, S. Figueiredo, R. L. Aguiar, H. Choo, "Distributed mobility management for the future mobile networks: A comprehensive analysis of key design options", IEEE Access, vol.5, pp. 11423-11436, 2017.

[12] A. Roy, J. Shin, N. Saxena, "Entropy-based location management in long-term evolution cellular systems", IET communications, vol.6, pp. 138-146, 2012.

[13] S. A. Hoseinitabatabei, A. Mohamed, M. Hassanpour, R. Tafazolli, "The Power of Mobility Prediction in Reducing Idle-State Signaling in Cellular Systems: A Revisit to 4G Mobility Management", IEEE Transactions on Wireless Communications, vol.19, pp. 3346-3360, 2020.

[14] Keyan Abdul Aziz Mutlaq and Hassanain Raheem Kareem and Ghida Yousif Abbass, "Modeling and Prediction LTE 4G NW According to MemoryAlgorithmof Long-Short Term", in International Journal of INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING, vol.11, pp. 172 – 178, 2023.