Signalling Cost Evaluation of Mobility Management Schemes for different Core Network Architectural Arrangements in 3GPP LTE/SAE

Meng Wang, Michael Georgiades and Rahim Tafazolli,
Centre for Communication System Research,
University of Surrey,
Guildford GU2 7XH, United Kingdom,
{M.Wang, M.Georgiades, R.Tafazolli}@surrey.ac.uk

***Abstract*— System Architecture Evolution (SAE) is one of the key challenges for the Long Term Evolution(LTE) of 3G research, charted by 3GPP, aiming to develop a new core network for improving the IP based packet-switched network performance. This paper considers different mobility management schemes for different SAE core network architectural options proposed in 3GPP, and evaluates their performance according to signalling cost by using analytical modelling based on the random walk model. Our analysis shows that the mobility management performance of the different options on both Idle and Active mode. Regarding the architecture, options 2 and 3 have the best trade-off in total signalling cost performance in both modes and signalling load on each network node for system scalability. For mobility management, Proxy MIPv6 has shown that it is able to effectively reduce the signalling cost for mobility management for all architecture options.**

Keywords —3GPP, Next generation network architecture, Mobility management, Mobile networks, Performance evaluation

I. INTRODUCTION

At present, the 3GPP [1] UMTS network employs the GPRS Tunnelling Protocol (GTP) for the mobility management across its radio access networks. All packet sessions are transported or relocated through the GPRS tunnel by the GPRS Mobility Management (GMM) procedure according to the Packet Data Protocol (PDP) context [2]. This mobility management functionality is optimized for the UMTS network to support both circuit-switched and packet-switched network architecture, but aimed mainly for the circuit-switched telephone service. However the high signalling overhead is not efficient for packet switched services. With the vast development in wireless communications, the user requirements for the wireless system are not so much related to circuit switched communication but mainly rise due from the great demanding high date rate pack-switch services for example music download, mobile TV. The evolution of the current 3GPP network aims to satisfy exactly this and meet these crescent requirements. Since 2004, 3GPP has begun charting the long-term evolution of 3G research for the new generation telecommunication systems referred to as 3GPP LTE. The aim is to evolve the 3GPP radio-access technology towards a high-data-rate, low-latency and packet-optimized radio-access technology, focusing on supporting services provided for the packet-switch domain. In parallel, the work item System Architecture Evolution (SAE) [3] has been approved to develop an Evolved Packet System (EPS) to accommodate the evolved radio access system. The SAE needs to support the migration of the existing 3GPP system and also the inter-working with non-3GPP radio access systems, such as WLAN, WiMAX, 3GPP2 etc. and will ultimately make the revolution leading toward All-IP wireless network [4], aiming to accommodate for the diverse demands of seamless mobility among heterogeneous wireless access services.

This paper provides an extended analysis as well as a performance evaluation of mobility management protocols for different architecture option proposals considered in 3GPP. The paper is organized as follow. Section 2 discusses the proposed architectures and mobility management protocol options, Section 3 presents our analytical modelling used for comparison, Section 4 deals with performance evaluation and discussion, and Section 5 concludes this paper.

II. ARCHITECTURE AND MOBILITY MANAGEMENT PROTOCOL OPTIONS

A. Packet Core Architecture Options

Developing new network architecture is recognized as one of the key goals of the SAE work item list. This evolves mainly around the decision of how to regroup the different functionalities into new logical entities with aim to best facilitate the 3GPP LTE access system. The EPS includes 3 logical entities namely Mobility Management Entity (MME), Serving Gateway (Serving GW) and Packet Data Network

978-1-4244-1645-5/08/$25.00 ©2008 IEEE

Gateway (PDN GW) [5][6]. The MME manages the functions of the Non-Access-Stratum (NAS) signalling and inter Core Network nodes mobility between 3GPP access networks, Idle mode UE tracking, reachability and authentication etc. The Serving GW is the gateway which terminates the interface towards the Evolved- Universal Terrestrial Radio Access Network (E-UTRAN) and includes the functions of the local mobility anchor point for inter eNB (evolved NodeB) handover, mobility anchoring for inter 3GPP mobility and packet routing and forwarding etc. The PDN GW is the gateway which terminates the interface towards the PDN and includes the functions of packet routing, forwarding and user plane anchor functions for mobility between 3GPP and non-3GPP access etc.

As discussed in [3][5][6][7], the different combinations of the logical entities result into four possible architecture options for the packet core network.

Fig. 1 presents the four options. As can be seen options 1 and 4 are the two extreme cases where the proposed logical entities are either fully separated or grouped together in one node. Option 2 is similar to SGSN/GGSN allocation in the UMTS network. For option 3, all the user plane functions are been grouped in one node, where as the other node manages the function for control plane. Different logical entities arrangements will result in different handling of the control signalling and user data packet routing. Integrating more logical entities in one component will result in less delay and signalling cost between different components, but increases the functional complexity per component.

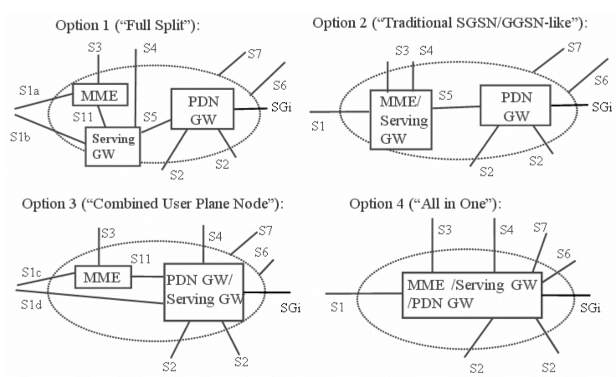


Fig. 1. Packet Core architecture options

B. Mobility Management Protocol Options

Regarding mobility management provisioning, three different mobility management protocol pairs (shown in Table 1) have been considered in 3GPP S2 WG and discussed in TR 23.882 [3]: The GTP plus GTP-U tunnelling based scheme (GTP+GTP-U), GTP plus MIPv6 scheme (GTP+MIP) and GTP plus PMIPv6 based scheme (GTP+PMIP).

Table 1. Mobility management scheme options

Interface / Scheme	S1	S5
GTP+GTP-U tunnelling	GTP	GTP-U tunnelling
GTP+MIPv6	GTP	MIPv6
GTP+PMIPv6	GTP	PMIPv6

As with the GTP tunnel between RNC and GGSN in UMTS [2], the data packets can be transferred through the GTP tunnel on the S1 interface between eNB and Serving GW (see Fig. 1). In the case of UMTS however, when a UE hands over from one GGSN to another, there is no mechanism to handover an active session seamlessly, as a new GTP-U tunnel between the UE and target GGSN needs to be re-established causing a break in the session. In order to enable the handover during an active session, the GTP+ GTP-U tunnelling scheme can setup the additional GTP-U tunnel between the Target Serving GW and the PDN GW using the packet forwarding mechanism to transfer the on-going session between Source Serving GW to Target Serving GW to enable handover between two Serving GWs.

For the GTP+MIPv6 based scheme, it uses GTP for the user data traffic tunneling between eNB and Serving GW, and adopts MIPv6 [8] for managing the user data traffic tunnel between the Serving GW and PDN GW and for enabling the inter-system handover during the active session. PMIPv6 [9] uses the proxy agent instead of HA for the routing update, so the UE does not need to additional support the function for MIPv6. As in the case of the GTP+MIPv6 based scheme, the GTP+PMIPv4 based scheme replaces the MIPv6 registration procedure with PMIP registration procedure to minimize the Mobile IP registration overhead.

Each packet core architectural arrangement with a different selection for a mobility management protocol gives rise to different mobility management performance. From the discussion of the 3GPP working groups, the proposals and contributions attempted to compare and identify the most suitable architecture and mobility management scheme for the LTE/SAE. However any resulting conclusions have mainly been based on discussions from previous experiences or by observation of the mobility management signalling flows.

III. Analytical Modeling

In order to give an in depth performance evaluation of the four packet core architecture options and three mobility management schemes, we propose an analytical model to calculate and compare the signalling cost for each arrangement of the packet core architecture options and mobility management scheme options. For analysis the two-dimensional hexagonal random walk model [9][11][12] has been adopted. As with the UMTS system the cells in an LTE system can be assumed to be configured as a hexagonal network with a cell having radio coverage of an eNB. The UE moves from one cell to another, and its movement is modelled based on the two-dimensional hexagonal random walk model. In this model, a hexagonal cell structure is modelled and the cells are classified in a 6-layer cluster shown in Fig. 2, which used to reduce the computational complexity and still remain the accuracy [12].

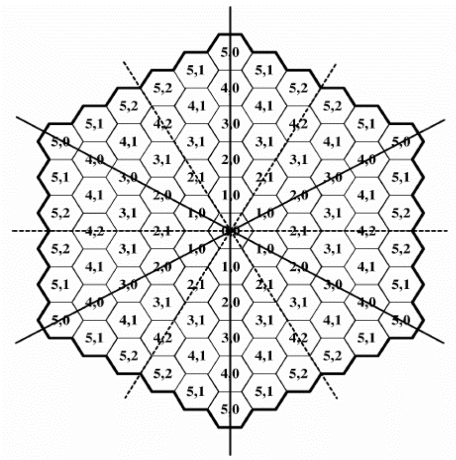


Fig. 2. Hexagonal cell structure and type classification in a 6-layer

We assume that a UE resides in a cell unit for a specified time period and then moves to any of the neighbouring cells with equal probability. Using this a one step transition matrix of this random walk can be derived by letting ***p(x,y),(x',y')*** be the one step probability from state (x, y) to (x', y'). The one step transition matrix of this random walk is matrix **P=**$\left(p(x,y),(x',y')\right)$, where:

$$\mathbf{P}=\begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1/6 & 1/3 & 1/6 & 1/3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1/6 & 0 & 1/3 & 1/6 & 1/3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1/3 & 1/3 & 0 & 0 & 1/3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1/6 & 0 & 0 & 1/3 & 1/6 & 1/3 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1/6 & 1/6 & 1/6 & 1/6 & 0 & 1/6 & 1/6 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1/6 & 0 & 0 & 1/3 & 0 & 1/6 & 1/3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1/6 & 1/6 & 1/6 & 0 & 1/6 & 0 & 1/6 & 1/6 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1/3 & 0 & 1/3 & 0 & 0 & 0 & 1/3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1/6 & 0 & 0 & 0 & 1/3 & 0 & 1/2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1/6 & 1/6 & 0 & 1/6 & 0 & 1/6 & 1/3 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1/6 & 1/6 & 0 & 1/6 & 1/6 & 1/3 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \quad (1)$$

Using the Chapman-Kolmogorov equation [9], the m-step transition matrix:

$$P^{(m)}=\begin{cases} P & , for\ m=1 \\ P\times P^{(m-1)} & , for\ m>1 \end{cases} \quad (2)$$

For r-layer cluster, define ***Pm(x,y),(r,j)*** as the probability that a UE initially resides at (x,y), then moves to (r-1, j) at m-1 step, and then moves out off the boundary at m-th step,

$$p^{(m)}{}_{(x,y),(r,j)}=\begin{cases} p_{(x,y),(r,0)} & , for\ m=1 \\ \sum_{j=0}^{N_{type}} p^{(m-1)}{}_{(x,y)}\times p_{(r-1,j),(r,0)} & , for\ m>1 \end{cases} \quad (3)$$

Probability of UE starts from arbitrary unit will leave the n-layer cluster at the mth movement:

$$\theta(n,m)=\sum_{i=1}^{N_{type}}\frac{n_i}{N_{unit}}p^{m}_{(x,y),(n,0)} \quad (4)$$

If a 6-layer cell structure as shown in Fig. 2. is considered, then:

$N_{unit}=3n^2-3n+1=91$, $N_{type}=3$

From the Markov Chain model, ***i*** is defined as the number of cell crossings since the last packet call ends, $\alpha_{i,j}$ is the state transition probability from state ***i*** to state ***j***, denoted by:

$$\alpha_{i,j}=\begin{cases} 1-\rho, & for\ j=i+1 \\ \rho, & for\ j=0 \\ 0, & otherwise \end{cases} \quad (5)$$

The probability that one or more packet calls arrive between two cell crossings can be calculated as follows:

$$\rho=\int_{t=0}^{\infty}(1-e^{-\lambda_s t})f_m(t)dt=1-f_m^*(\lambda_s)=1-\left(\frac{\gamma\lambda_m}{\lambda_s+\gamma\lambda_m}\right)^{\gamma}\overset{\gamma=1}{=}\frac{\lambda_s}{\lambda_s+\lambda_m} \quad (6)$$

From the probability above, we can model the signalling cost of two mobility states in an LTE system: LTE_ACTIVE where the network directs UE to the serving cell and the UE is ready to perform Uplink/Downlink transport with very limited access delay, and LTE_IDLE where the UE in a low power consumption state, could be tracked in Tracking Area and be able to transit to LTE_ACTIVE about 100ms.

For the LTE_ACTIVE mode, the signalling procedure costs are defined: Total Handover Cost is the handover cost when an on-going application session when UE moves across the cells; Total Session Activation Cost is the cost to setup the transport tunnel for new arrived session. We define ***H*** as the Handover Cost for one handover process, ***A*** as the session activation cost for one session activation process. Handover cost during UE stays in state ***i*** of Markov Chain,

$$c_h(i)=\theta(n,i)\times H \quad (7)$$

Average handover cost per state transition is,

$$c_h(i)'=\sum_{i=0}^{\infty}p_i c_h(i)=\rho\sum_{i=0}^{\infty}(1-\rho_i)^i c_h(i) \quad (8)$$

Average handover cost per unit time is,

$$C_h=\lambda_m c_h' \quad (9)$$

Assuming ***v*** is the number of packet call arrivals between two cell crossing, then $v=\lambda_s/\lambda_m$. Session activation cost during UE stay in state ***i*** of Markov Chain,

$$c_a(i)=v\times N_{unit}\times A \quad (10)$$

The average session activation cost per state transition is,

$$c_a(i)'=\sum_{i=0}^{\infty}p_i c_a(i)=\rho\sum_{i=0}^{\infty}(1-\rho_i)^i c_a(i) \quad (11)$$

The average session activation cost per unit time is,

$$C_a=\lambda_m c_a' \quad (12)$$

Total signalling cost for LTE_ACTIVE mode is

$$C_T=C_h+C_a \quad (13)$$

In the LTE_IDLE mode, the signalling cost includes: Total

Location Update Cost when the Core network update the UE location as UE moving across the Tracking Area (TA), and the Total Paging Cost when the Core network pages all the cells in a TA to track the position of the UE. From the same equation above, but with the value for location update cost $\boldsymbol{L}$ and paging cost $\boldsymbol{P}$ replace the $\boldsymbol{H}$ and $\boldsymbol{A}$, we can have the total signalling cost for the LTE_IDEL mode,

$$C_T = C_l + C_p \tag{14}$$

If we replace the total signalling cost of the whole network with the signaling cost on certain network nodes, we can obtain the signaling cost on each network node to evaluate the scalability for the network.

IV. PERFOMANCE EVALUATION

Based on the analytical model described in Section III, this section conducts a performance evaluation based on the analytical modelling results for both LTE_ACTIVE and LTE_IDLE mode. For calculating the signalling cost, we follow the signalling exchange chart drafted in 3GPP standard [3][5][6]. We use the signalling message sizes to sum up the total signalling cost and signalling load on each core network nodes. For the signalling message sizes we use real 3G operational values based on consultation from Motorola Ltd.

A. *Performance evaluation for the LTE_ACTIVE mode*

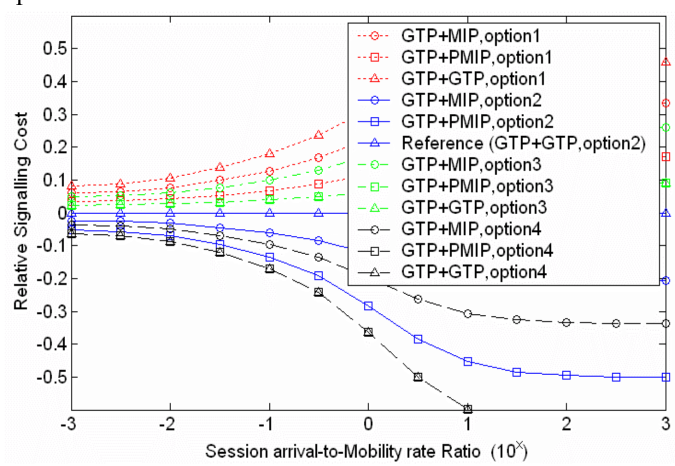
Using the total handover cost and the total session activation cost, we could obtain the total signalling cost per packet call arrival $\boldsymbol{C_T}$, based on the Session arrival to Mobility rate (SMR) λ_s/λ_m, where λ_s is the packet Session arrival rate and λ_m is the Mobility rate. High SMR means that the packet session arrival rate is higher than the mobility rate which in practice it has high application session arrival before a serving cell handover. Low SMR means the packet session arrival rate is lower than the mobility rate which in practice means more cell handovers per single session. For y-axis, it shows the relative signalling cost which all the different architecture and mobility management scheme arrangements comparing with the reference arrangement option 2 using GTP+GTP, which is the most similar to the GTP based UMTS core network architecture.

Fig. 3(a) shows the relative total signalling cost for all different architectures and mobility management scheme arrangements in the LTE_ACTIVE mode case. For options 1 and 2, the GTP+GTP scheme introduces higher signalling cost than the GTP+MIP, TP+PMIP schemes. This is because GTP+GTP involving more signalling procedure for the user-plane tunnel handover compared with MIP and PMIP. In options 3 and 4, GTP+GTP and GTP+PMIP have same signalling cost, and less than GTP+MIP. In options 3 and 4, the user-plane tunnelling process in handled by the Serving GW and PDN GW entities, same applies to all schemes except GTP+MIP where it requires additional signalling to eNB for the mobile IP registration process.

Fig. 3(b) shows the relative signalling traffic load on the radio link between eNB and UE. There is no difference between the different architecture options, but when comparing mobility management schemes GTP+MIP has more signalling traffic than any other. This is again due to the mobile IP registration process which transfers more signalling on the radio link.

The PDN GW node in options 1 and 2, the Serving GW/PDN GW combine node in option3 and MME/Serving GW/PDN GW combine node in option 4 consist the gateway function of anchoring the user plane traffic tunnel for core network. These anchor nodes handle most of the signalling procedures for tunnelling control and is the most loaded node regarding user traffic in the core network, so it has a major impact on the core network performance. Fig. 4(c) shows the relative signalling traffic load comparison at the anchor node. For different architectures: the S5 interface of options 1 and 2 have the same signalling traffic cost; for options 3 and 4 the signalling traffic at the anchor node also includes the signalling to the eNB. For options 1, 2 and 3, the tunnelling cost for handover and packet session arrival is the same, so the relative signalling cost remains constant. For option 4, the control plane signalling for eNB is also transferred to the anchor, so with a higher mobility rate the relative signalling cost becomes much higher. For different mobility schemes, GTP+GTP has more signalling cost at the anchor node than GTP+MIP and GTP+PMIP due to the tunnelling update procedure cost .

Fig. 4(d) shows the relative signalling traffic load at the MME node, which is the most heavily loaded regarding control-plane traffic. S1a and S11 in option 1 and S1c and S11 in option 3 have the same amount of signalling traffic for MME control over eNB and Serving GW, and the differences in cost for mobility management schemes are for controlling the user-plane tunnel between Serving GW and PDN GW, so options 1 and 3 with different mobility management schemes have the same signalling traffic load at the MME. In options 2 and 4 the MME and Serving GW are co-located which results to less traffic at the MME for options 2 and 4 as compared to options 1 and 3.



(a)

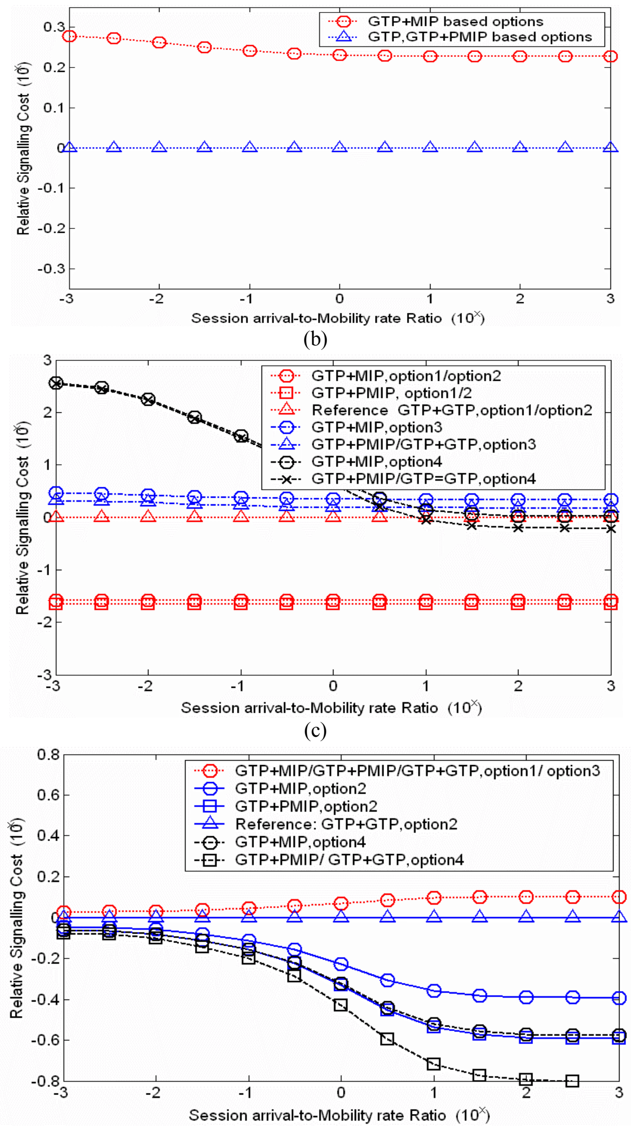


Fig. 3. Relative signalling cost for LTE_ACTIVE mode

B. *Performance evaluation for the LTE_IDLE mode*

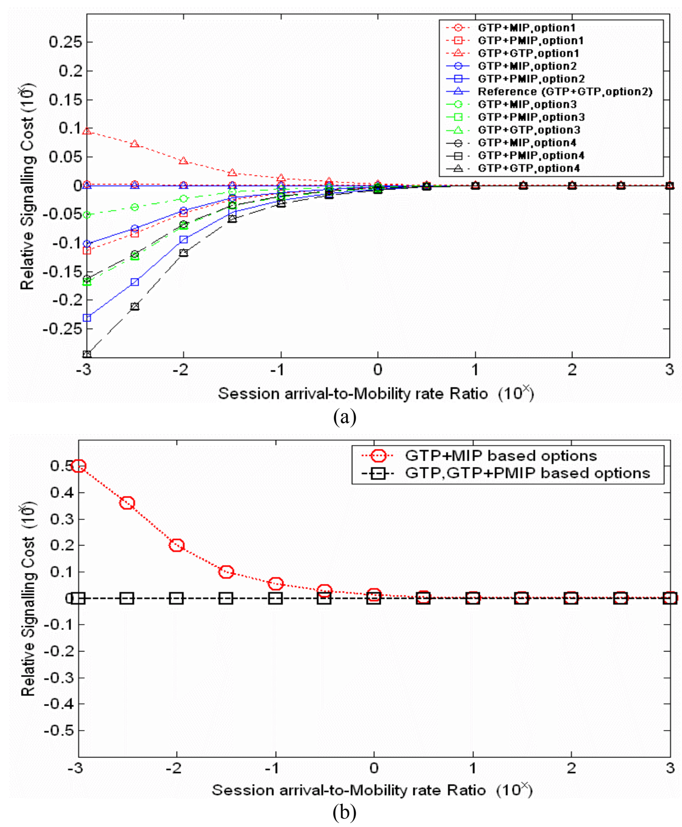
For the LTE_IDLE mode, the results show the relative signalling cost against the SMR, where the reference is option 2 using GTP+GTP. High SMR implies that the paging rate is higher than the location update rate, which means that the UE a higher probability to page the UE in LTE_IDLE mode for setup the connection when UE resides in a TA; Low SMR means that the location update rate is higher than the paging rate which means that the network has more location update process in LTE_IDLE mode when moving across different TAs. We assume that all the different mobility management schemes use the same mechanism for the paging process, so the signalling cost is the same when the paging rate is much higher than the location update rate, in the case where the paging process is more dominant than the signalling traffic.

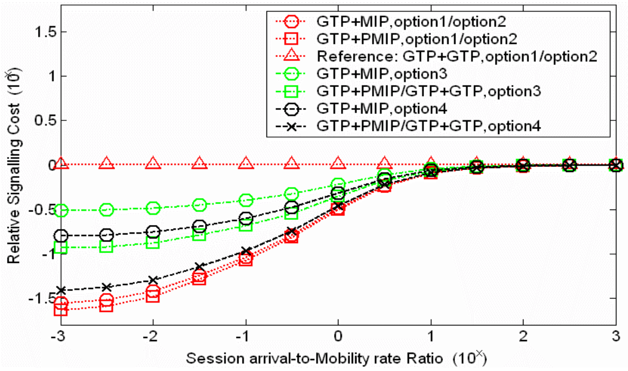
Fig. 4(a) shows the relative total signalling cost for all the different architectures and mobility management schemes arrangements for the LTE_IDLE mode. In both architectural options 1 and 2, GTP+GTP has more signalling traffic than GTP+MIP and GTP+PMIP. For the same mobility management scheme, option 1 faces more signalling traffic than option 2. For options 3 and 4, because the Serving GW and PDN GW are combined under one node, the user-plane update procedure is the same for GTP+GTP and GTP+PMIP. However for the GTP+MIP the user-plane update procedure is higher due to the extra mobile IP registration process between eNB and GW.

Fig. 4(b) shows the relative signalling traffic load across the radio link between eNB and UE. Same as in the LTE_ACTIVE mode case, it can be seen that GTP+MIP has more signalling traffic than the other two due to the mobile IP registration process.

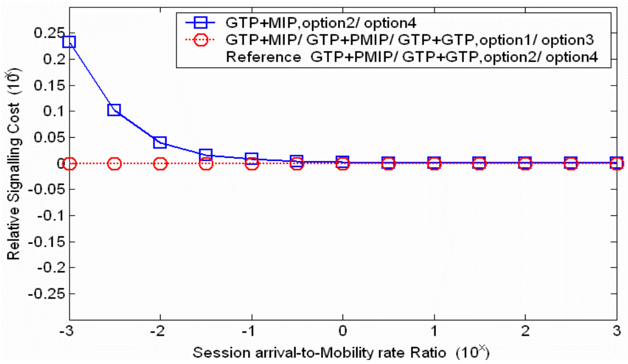
Fig. 4(c) shows the relative signalling traffic load comparison at the anchor node. For options 3 and 4, GTP+MIP has more signalling load due to the mobile IP registration signalling. Options 1 and 2, have the same signalling transfer through the S5 interface. GTP+GTP involves more signalling for location update at the anchor node which means more signalling load than the other two mobility management schemes.

Fig. 4(d) shows the relative signalling traffic load comparison for MME node for LTE_IDLE mode. For the GTP+GTP and GTP+PMIP, the signalling traffic for different architectural options is the same; as they have the same signalling at the MME. GTP+MIP needs the mobile IP registration procedure transfer through MME for options 2 and 4, therefore the two arrangements have more signalling load at the MME.





(c)



(d)

Fig. 4. Relative signalling cost for LTE_IDLE mode

V. CONCLUTION AND FUTURE WORKS

This paper gives an introduction of the different SAE core network architectures and mobility management schemes that have been considered in the 3GPP SAE working groups. An analytical model based on the two-dimensional hexagonal random walk model is proposed for comparing the mobility management performance for both LTE_ACTIVE mode and LTE_IDLE mode. The results give a detailed performance evaluation of all mobility management/architectural option arrangements considered (4 logical architectural options and 3 mobility management schemes). Our performance analysis indicates that: option 4 has the best total signalling performance but introduces heavy signalling load on the network node, but has the worst performance of the load on network load for scalability; option 1 has less signalling load on network nodes comparing with other options; option 2 and option 3 have the best trade-off in total signalling cost and signalling load on each network node. For each of the options, Proxy MIPv6 has great advantage on reducing the signalling cost of mobility management. Option 2 or option 3 with the mobility management scheme Proxy MIPv6 could be the recommend arrangement for SAE core network. For future work, the handover delay performance for user-plane interruption has been carrying on, as the handover delay is critical in providing a satisfactory user experience. This study will ultimately lead to optimise the mobility management for a selected SAE core network architecture.

ACKNOWLEDGMENT

This work has been carried out at the Centre of Communication Systems Research and supported by Motorola Ltd. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements either expressed or implied, of Motorola Ltd. The authors would like to especially thank Dr. Howard Benn, Dr. Howard Thomas and Dr. Luis Lopes from Motorola Ltd. for their valuable comments.

REFERENCES

[1] Third Generation Partnership Project (3GPP), http://www.3gpp.org/
[2] 3GPP TS 23.060, “General Packet Radio Service (GPRS); Service description; Stage 2”, Rel. 7, v 7.2.0, Oct. 2006.
[3] 3GPP TR 23.882, “3GPP system architecture evolution (SAE): Report on technical options and conclusions, Rel. 7, v1.9.0, April 2007.
[4] 3GPP TR 22.978, “All-IP network (AIPN) feasibility study”, Rel. 7, v7.1.0, June 2005.
[5] 3GPP TS 23.401, “General Packet Radio Service (GPRS) enhancements for Long Term Evolution (LTE) access”, Rel. 8, v1.1.1, Aug. 2007.
[6] 3GPP TS 23.402, “3GPP System Architecture Evolution (SAE): Architecture enhancements for non-3GPP accesses”, Rel. 8, v1.2.1, Aug. 2007.
[7] 3GPP S2-060820, “Grouping of Core Network Entities”, S2#51meeting, Feb. 2006.
[8] D. Johnson, C. Perkins, and J. Arkko, "Mobility Support in IPv6". IETF, RFC 3775, June 2004.
[9] S. Gundavelli et. Al.: "Proxy Mobile IPv6", Internet Draft, draft-sgundave-mip6-proxymip6-00, October 2006.
[10] Akyildiz, I. F., Ho, S. M., and Lin, Y.-B. Movement-Based Location Update and Selective Paging for PCS Networks. IEEE/ACM Transactions on Networking, 4(4):629-638, August 1996.
[11] Akyildiz, I. F., Lin, Y.-B., Lai, W.-R., and Chen, R.-J. A New Random Walk Model for PCS Networks. IEEE Journal on Selected Areas in Communications, 18(7):1254-1260, July 2000.
[12] Yang, S.-R. and Lin, Y.-B. Performance Evaluation of Location Management in UMTS. IEEE Transactions on Vehicular Technology, 52(6):1603-1615, November 2003.