Wireless Netw (2017) 23:1645–1654
DOI 10.1007/s11276-016-1374-0

Impact of mobile instant messaging applications on signaling load and UE energy consumption

Yunjian Jia[1] · Yu Zhang[1] · Liang Liang[1] · Weiyang Xu[1] · Sheng Zhou[2]

Published online: 25 October 2016

Abstract Mobile Instant Messaging (MIM) applications transmit not only user-triggered messages (UTMs), but also keep-alive messages (KAMs) via radio access network, which induces heavy burden in control plane channel and wastes user equipment (UE) energy consumption. In this paper, we deduce the joint distribution of KAM period and UTM mean interval from the MIM application traffic characteristics. Correlating the joint distribution with radio resource control (RRC) state machine in LTE networks, we derive two analytical expressions for the control plane signaling load and UE energy consumption respectively. Then, the variation of signaling load and energy usage is demonstrated with different settings of RRC release timer, KAM period and UTM mean interval. The analysis indicates that KAM period is the upper bound of RRC release timer when reducing the signaling load. Besides, five times of UTM mean interval is the upper bound of KAM period when reducing the UE energy consumption and signaling load. These results can guide both network operators and MIM application developers to properly set control parameters for balancing the signaling load and UE energy consumption.

✉ Yunjian Jia
yunjian@cqu.edu.cn

[1] Key Laboratory of Dependable Service Computing in Cyber Physical Society (Chongqing University), Ministry of Education; College of Communication Engineering, Chongqing University, Chongqing 400044, China

[2] Tsinghua National Laboratory for Information Science and Technology, Department of Electronic Engineering, Tsinghua University, Beijing 100084, China

Keywords Mobile instant messaging application · Radio resource control state machine · Signaling load · UE energy consumption

1 Introduction

Mobile instant messaging (MIM) applications are widely adopted in user equipments (UEs) [1–3]. They differ from traditional telecom services, such as voice and short message, in that they provide an always-on connectivity [4, 5]. Therefore, they periodically and autonomously send small packets to maintain the Internet Protocol (IP) connection with application servers even when the applications are out of user interaction. These small packets are so-called keep-alive messages (KAM). However, the transmission of KAM in radio access network (RAN) leads to frequent transition between different radio resource control (RRC) states, thus induces heavy burden in control plane channel [6–9] and wastes UE energy consumption [9–12].

To maximum the spectrum efficiency and the energy efficiency, authors in [13] proposed a resource allocation solution based on non-cooperative game theory for D2D communication. Besides, authors in [14] developed a QoE aware spectrum efficiency and QoE aware energy efficiency based mobile association and resource allocation scheme in a heterogeneous wireless network. However, these solutions are not suitable for MIM applications since the traffic pattern and resource requirements between MIM application and D2D communication are not the same. Therefore, we mainly focus on the MIM traffic's requirements on radio resources and corresponding impacts on network and user in this paper.

RRC states are controlled by RRC release timer, which is set by the evolved NodeB (eNB). Authors in [15–18]

found that longer RRC release timer brings lower control plane signaling load but higher energy consumption and vice versa. Hereinafter, signaling load is used to represent control plane signaling load for simplification. Taking the application traffic into consideration, authors in [19, 20] found that, longer RRC release timer brings lower signaling load only when packet arrival rate is relatively low. If the packet arrival rate is high enough, increasing RRC release timer will end in higher signaling load. As to the UE energy consumption, shorter RRC release timer always consumes less energy regardless of the packet arrival rate. Besides, authors in [21] found that RRC release timer influences signaling load and energy consumption only when the maximum inter-arrival time between two consecutive packets is less than the RRC release timer. However, various packets within the application traffic, including KAM and other user-triggered messages (UTM) were treated with no difference in the above works. Consequently, how the KAM influences the signaling load and energy usage was not formulated, although KAM transmission is the main reason for the heavy burden in control plane channel and unnecessary energy consumption of UE.

In this paper, we focus on MIM applications in LTE networks. Firstly, we deduce the joint distribution of KAM period and UTM mean interval based on the MIM traffic characteristics. Correlating the joint distribution with RRC state machine, we derive two analytical expressions for signaling load and UE energy consumption respectively. Then, the variation of signaling load and energy usage is demonstrated with different settings of RRC release timer, KAM period and UTM mean interval. The rest of this paper is organized as follows: Sect. 2 describes the system model, which includes the MIM application traffic model and its correlation with the RRC state machine. In Sect. 3, we characterize the signaling load and UE energy consumption with two analytical expressions. We then study the influence of RRC release timer and different messages intervals on the performance of signaling load and UE energy consumption respectively in Sect. 4, optimization suggestions are provided simultaneously. Finally, Sect. 5 contains our conclusion, followed by the future works in Sect. 6.

2 System model

To analyze the impact of MIM applications on signaling load and UE energy consumption, MIM traffic is modeled according to 3GPP Technical Report [22]. Correlating the MIM traffic with RRC state machine, we can depict the RRC state transition diagram accordingly, which can be used in modeling the signaling load and UE energy consumption.

2.1 Traffic model of MIM applications

Figure 1 illustrates the traffic model of MIM application, where t_m^n, t_{KA}, t_m and t_R represent the UTM interval, the KAM period, the message interval that joint two kinds of messages and the RRC release timer, respectively. In the beginning, the MIM application logs in and receives the OK message from the server. Then, the MIM application can transmit and receive messages via RAN. Acknowledgement (ACK) message is replied from the receiving side. Note that, the LOGIN, OK, TX, ACK, RX and LOGOUT messages are so called UTM messages, which are denoted by solid arrows in Fig. 1. Specifically, the blue upward arrows denote the uplink UTM messages, and the red downward arrows denote the downlink UTM messages. Generally, the length of these messages is in the range of tens of bytes to few kilobytes, thus the transmission time of these messages is negligible small under the context of high transmission rate in LTE networks. Besides, the reaction time of network or UE is in the range of tens of milliseconds, which is also negligible small compared to the message interval. Therefore, we only consider the message interval in this paper. The UTM interval, i.e., t_m^n, is exponentially distributed with mean value of $1/\lambda$ according to [22]. Specifically, $1/\lambda$ is decided by user activity level. The generation of KAM is controlled by a timer, called KAM period, i.e., t_{KA}, which is set by application developer during the development stage. Dotted arrows are used to represent the KAM messages in Fig. 1. Similarly, the blue upward arrows denote the uplink KAM messages, and the red downward arrows denote the downlink KAM messages. Whenever t_m^n surpasses t_{KA}, the KAM is generated autonomously and periodically until next UTM arrives, so that the long-lived IP connection can be maintained. We use t_m to denote the message interval that joint KAM period and UTM interval. Therefore, t_m is the remainder of t_m^n divided by t_{KA}:

$$\{t_m\} = \begin{cases} \{t_m^n\}, & 0 \le t_m^n \le t_{KA} \\ \left\{ t_m^n - i \cdot t_{KA}, \underbrace{t_{KA}, \ldots, t_{KA}}_{i} \right\}, & t_m^n > t_{KA} \end{cases} \tag{1}$$

where $i = [\frac{t_m^n}{t_{KA}}]$, and $[\cdot]$ is the floor function.

Consequently, t_m is head-truncated distributed. Hence, typical exponential distribution cannot describe t_m anymore. Fortunately, the probability density function (PDF) of t_m when t_m is in the range of $[0, t_{KA})$ is still the same with t_m^n, and mathematical derivation can be found in Appendix 1. Thus, the PDF of t_m can be derived as follows:

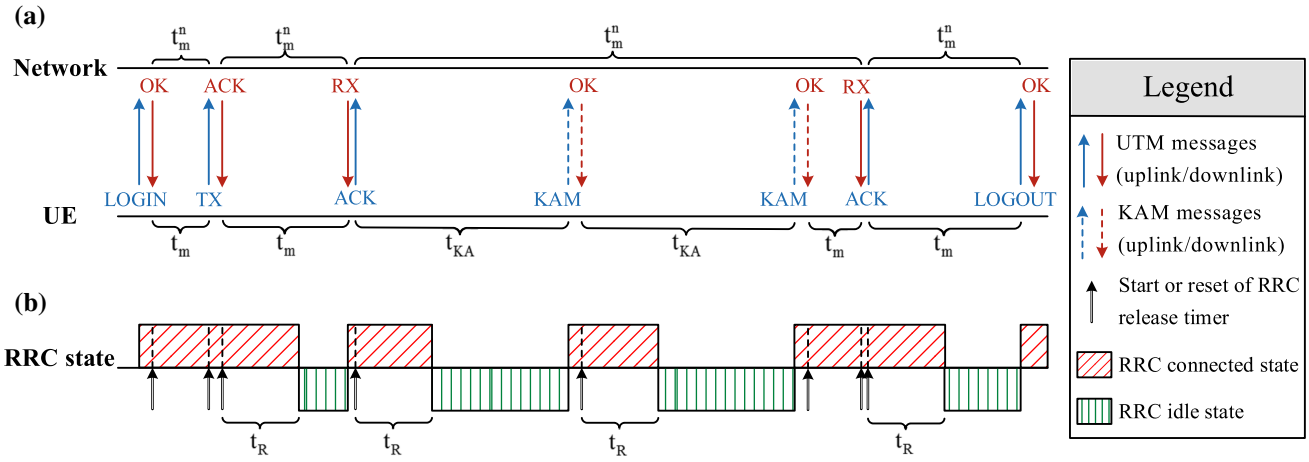


Fig. 1 MIM traffic model and RRC state transition diagram. **a** MIM traffic model. **b** RRC state transition diagram

$$f(t_m) = \begin{cases} \lambda e^{-\lambda t_m}, & 0 \le t_m < t_{KA} \\ \int_{t_{KA}}^{+\infty} \lambda e^{-\lambda t_m} dt_m, & t_m = t_{KA} \\ 0 & t_m > t_{KA}. \end{cases} \quad (2)$$

Furthermore, the mean message interval of joint messages, i.e., $\overline{t_m}$, can be calculated as follows:

$$\overline{t_m} = \int_0^{+\infty} t_m \cdot f(t_m) dt_m = \frac{1}{\lambda} - \left(\lambda \cdot t_{KA} + \frac{1}{\lambda}\right) e^{-\lambda t_{KA}}. \quad (3)$$

Since message length makes no sense in RRC state transitions, we ignore the message length in the traffic model for simplification.

2.2 RRC state machine in LTE networks

In LTE networks, there are two RRC states including RRC connected state and RRC idle state [23]. A UE can only transmit and receive data in the RRC connected state. For a UE in RRC idle state, RRC connection establishment is needed when a new message arrives. After data transmission is completed, RRC release timer starts counting. If there is no message activity until RRC release timer expires, eNB will switch the UE from RRC connected state to RRC idle state. Thus, message interval and RRC release timer are the parameters that influence RRC state transition. We use t_R to denote RRC release timer in this paper. RRC state transition diagram can be depicted on the basis of MIM traffic, as shown in Fig. 1. The start/reset of RRC release timer is denoted by the black solid upward arrows in Fig. 1. Besides, the box fixed with red backslash represents the RRC connected state, and the box fixed with green vertical stripe represents the RRC idle state. The radio power is 290 mA in RRC connected state and 4 mA in RRC idle state, which is analyzed by Agilent power analyzer.

3 Performance metrics

Based on the MIM traffic and RRC state machine described in Sect. 2, the RRC state machine transition diagram can be depicted accordingly as shown in Fig. 1(b). By correlating message interval with RRC release timer in the transition diagram, the signaling load, and UE energy consumption can be derived accordingly.

3.1 System signaling load

Given that the heavy signaling load caused by MIM applications mainly results in the RRC state transition, we only take state transition related signaling into consideration. A complete RRC connection establishment and corresponding connection release process represent a whole RRC state transition. We denote the number of RRC state transitions by N_{trans}.

For a UE in RRC connected state, if next message is generated within RRC release timer, i.e., $t_m \le t_R$, then the UE will stay in RRC connected state without a state transition; but if next message is generated after the expiration of RRC release timer, i.e., $t_m > t_R$, the UE will switch to RRC idle state, and then get back to RRC connected state. Therefore, the UE transition probability P_t can be calculated as follows:

$$P_t = \Pr\{t_m > x\} = \begin{cases} e^{-\lambda x}, & 0 \le x < t_{KA} \\ 0, & x \ge t_{KA} \end{cases} \quad (4)$$

where x denotes the value of t_R.

Since N_{trans} is linearly proportional to P_t and the number of messages, N_{trans} can be calculated as follows:

$$N_{trans} = T_0 \cdot R_m \cdot P_t, \tag{5}$$

where T_0 stands for the unit time, and R_m represents the arrival rate of messages. Since the transition time and reaction time of messages are ignored in this paper as detailed in Sect. 2, R_m can be calculated with the help of mean message interval:

$$R_m = 1/\overline{t_m}, \tag{6}$$

Since a whole state transition consumes eight signaling messages in RAN side as shown in Fig. 2, the average signaling load per unit time, i.e. S can be calculated as follows:

$$S = 8N_{trans} = \frac{8T_0 \cdot \Pr\{t_m > x\}}{\overline{t_m}}, \tag{7}$$

Substitute (3), (4) into (7), the full expression of S can be rewritten as follows:

$$S = \begin{cases} \dfrac{8T_0 \cdot e^{-\lambda x}}{\dfrac{1}{\lambda} - \left(\lambda \cdot t_{KA} + \dfrac{1}{\lambda}\right) e^{-\lambda t_{KA}}}, & 0 \le x < t_{KA} \\ 0, & x \ge t_{KA} \end{cases} \tag{8}$$

When $t_{KA} > 5\frac{1}{\lambda}$, $e^{-\lambda t_{KA}}$ is smaller than 0.01. Since $1/\lambda$ is generally in the range of few tens of seconds to few hundreds of seconds, $1/\lambda > 100(\lambda \cdot t_{KA} + \frac{1}{\lambda})e^{-\lambda t_{KA}}$. Therefore, $(\lambda \cdot t_{KA} + \frac{1}{\lambda})e^{-\lambda t_{KA}}$ is negligibly small. Then (8) can be further simplified as follows:

$$\text{when } t_{KA} > 5\frac{1}{\lambda}, S = \begin{cases} \lambda \cdot 8T_0 \cdot e^{-\lambda x}, & 0 \le x < t_{KA} \\ 0, & x \ge t_{KA} \end{cases} \tag{9}$$

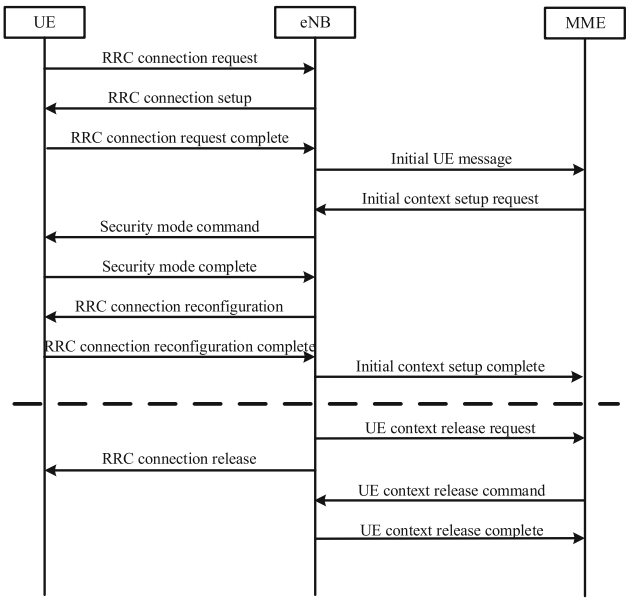


Fig. 2 Connection establishment and release procedures

3.2 UE energy consumption

The energy consumption for a UE mainly depends on the time it spends in RRC connected state, i.e., T_c. For a specific message of a UE in RRC connected state, the message interval t_m is known. If $t_m \le t_R$, then no state transition will occur, and t_m will determine how long the UE spends in RRC connected state caused by this message. On the contrary, if $t_m > t_R$, then the UE will transfer to RRC idle state after the expiration of RRC release timer. Therefore, t_R will determine how long the UE spends in RRC connected state. Thus, for a chain of messages, the average time for a UE to remain in RRC connected state per message, i.e., $\overline{t_c}$ is determined by:

$$\overline{t_c} = \begin{cases} \Pr\{t_m \le x\} \cdot h(x) + \Pr\{t_m > x\} \cdot x, & 0 \le x < t_{KA} \\ \overline{t_m}, & x \ge t_{KA} \end{cases} \tag{10}$$

where $h(x)$ denotes the mean value of t_m when $0 \le t_m \le x (0 \le x < t_{KA})$. The full expression of $h(x)$ is as follows:

$$h(x) = \frac{\left(x + \frac{1}{\lambda}\right) \cdot e^{-\lambda x} - \frac{1}{\lambda}}{e^{-\lambda x} - 1}. \tag{11}$$

The calculation of $h(x)$ can be found in Appendix 2. Therefore, the whole time UE spends in RRC connected state T_c can be calculated as follows:

$$T_c = T_0 \cdot R_m \cdot \overline{t_c}, \tag{12}$$

Since the time of every RRC state transition is smaller than 1 ms and the power of RRC state transition is relatively low, we ignore the energy consumed in the RRC state transition. Therefore, the average energy consumption per unit time can be calculated as follows:

$$E = 290T_c + 4(T_0 - T_c) = 284T_c + 4T_0 = 284T_0 \cdot \frac{\overline{t_c}}{\overline{t_m}} + 4T_0. \tag{13}$$

Substitute (3) and (10) into (13), the full expression of E can be rewritten as follows:

$$E = \begin{cases} \dfrac{284T_0 \cdot \left(\dfrac{1}{\lambda} - \dfrac{1}{\lambda} \cdot e^{-\lambda \cdot x}\right)}{\dfrac{1}{\lambda} - \left(\lambda \cdot t_{KA} + \dfrac{1}{\lambda}\right) e^{-\lambda t_{KA}}} + 4T_0, & x < t_{KA} \\ 288T_0, & x \ge t_{KA} \end{cases} \tag{14}$$

As detailed before, when $t_{KA} > 5\frac{1}{\lambda}$, $(\lambda \cdot t_{KA} + \frac{1}{\lambda})e^{-\lambda t_{KA}}$ is negligibly small. Then (14) can be further simplified as follows:

$$\text{when } t_{KA} > 5\frac{1}{\lambda}, E = \begin{cases} 284T_0(1 - e^{-\lambda \cdot x}) + 4T_0, & 0 \le x < t_{KA} \\ 288T_0, & x \ge t_{KA} \end{cases} \quad (15)$$

4 Numerical analysis

In this section, we study the variation of signaling load and UE energy consumption caused by MIM applications when RRC release timer, KAM period and UTM mean interval varies. We use the traffic model described in Sect. 2 to simulate the MIM traffic and calculate S and E based on (8) and (14). Three scenarios are considered in this simulation. The first one is a fixed MIM application with different user activity levels. The t_{KA} of this application is set to 120s [24], which is widely used in MIM applications. To represent different user activity levels, $1/\lambda$ is set to 20s, 40s, 80s and 160s, respectively. In the second scenario, the user activity is fixed with different MIM applications. Therefore, $1/\lambda$ is fixed to 40s, while t_{KA} is set to 30s, 60s, 180s and 600s, respectively. In the last scenario, a UE has multiple MIM applications. We assume the number of MIM applications is 5. The KAM periods of the 5 applications are denoted by $t_{KAi}, i \in [1, 2, \ldots, 5]$, while the UTM mean intervals of the 5 applications are denoted by $t^n_{mi}, i \in [1, 2, \ldots, 5]$. In this scenario, the minimum t_{KAi} is set to 60s, 120s and 240s, respectively, while the average t^n_{mi} is set to 120s and 60s, respectively. Detailed simulation parameters and assumptions are provided in Table 1, which keeps consistent with the parameters in [24].

4.1 System signaling load

The average signaling load per unit time is given in Fig. 3. For the first two scenarios, when t_R is smaller than t_{KA}, the signaling load decreases while t_R gets longer as shown in Fig. 3(a, b). The detailed reason can be found in previous works [19, 20]. Moreover, the decreasing rate for the application with lower $1/\lambda$ is higher than the application with longer $1/\lambda$. In other words, longer t_R is effective in decreasing the signaling load especially when $1/\lambda$ is relatively low. Besides, the signaling load decreases while t_{KA} gets longer as shown in Fig. 3(b). However, when t_{KA} increases to about five times of $1/\lambda$, it makes no difference to go on to increase t_{KA}. This can be explained by (9). Thus, we can draw the conclusion that t_{KA} is upper bounded in the course of reducing signaling load. The upper bound is about five times of UTM mean interval. This can be used for application operator to configure the KAM period. Once t_R is larger than t_{KA}, the signaling load turns to 0 immediately as shown in Fig. 3(a, b). This is due to the fact that t_m is always shorter than t_R in this case. The UE keeps in RRC connected state, and no state transition occurs no

Table 1 Simulation parameters

Scenario I	
Description	Fixed MIM application with different user activity
Simulation time (h)	5
Unit time (h)	1
RRC release timer (s)	1, 4, 16, 64, 256, 1024
KAM period (s)	120
UTM mean interval (s)	20, 40, 80, 160
Scenario II	
Description	Fixed user activity with different MIM applications
Simulation time (h)	5
Unit time (h)	1
RRC release timer (s)	1, 4, 16, 64, 256, 1024
KAM period (s)	30, 60, 180, 600
UTM mean interval (s)	40
Scenario III	
Description	A UE with multiple MIM applications
Simulation time (h)	5
Unit time (h)	1
RRC release timer (s)	1, 4, 16, 60, 120, 240, 1024
KAM period (s)	[300, 240, 180, 120, 60]; [360, 300, 240, 180, 120]; [480, 420, 360, 300, 240]
UTM mean interval (s)	[90, 105, 120, 135, 150]; [30, 45, 60, 75, 90]

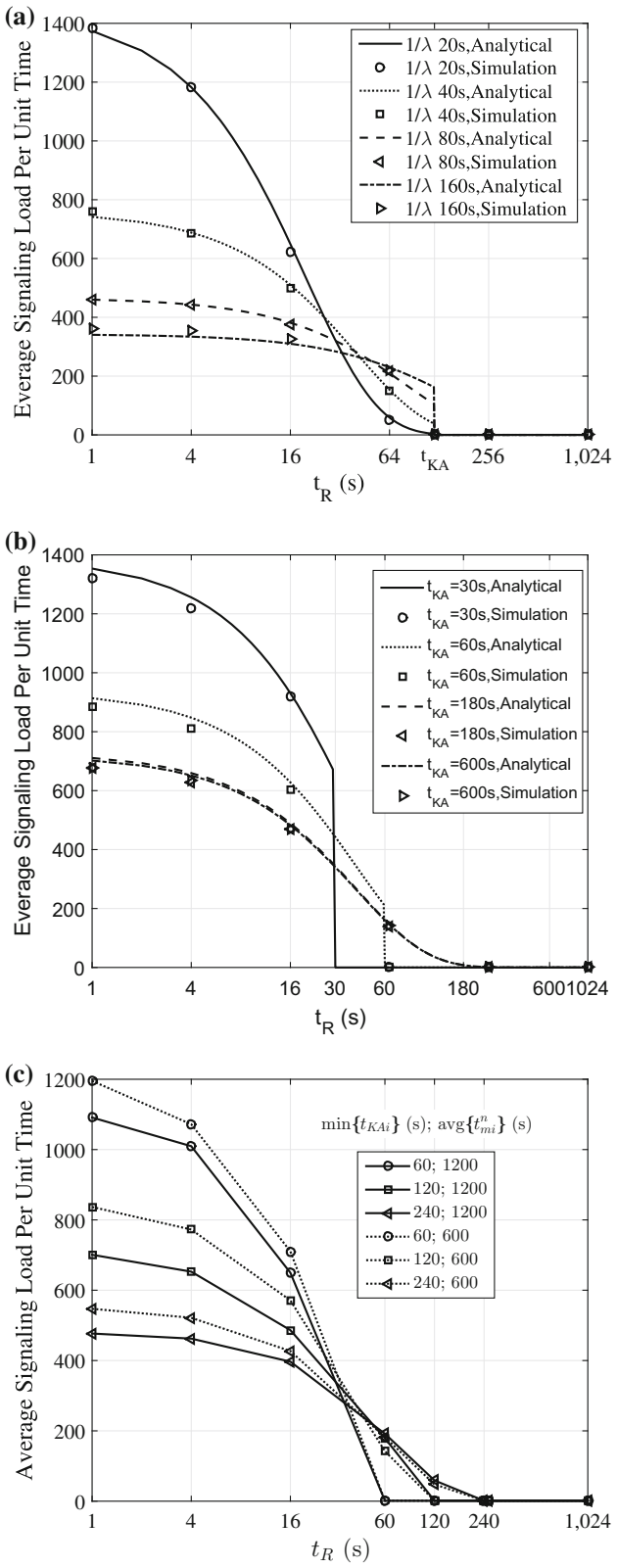


Fig. 3 System signaling load per unit time. **a** S (Scenario I). **b** S (Scenario II). **c** S (Scenario III)

matter what the value of $1/\lambda$ and t_R are. For the UE with multiple MIM applications, when t_R is smaller than the $\min\{t_{KAi}\}$, longer t_R is still effective in decreasing the signaling load especially when the $\text{avg}\{t^n_{mi}\}$ is relatively low. Once t_R surpasses the $\min\{t_{KAi}\}$, the signaling load turns to 0 immediately. This is because the maximum t_m of the UE is limited to the $\min\{t_{KAi}\}$ with the superposition of multiple applications. Therefore, we can draw the conclusion that t_R is upper bounded in the course of reducing signaling load, and the upper bound is t_{KA} for the UE with single MIM application and $\min\{t_{KAi}\}$ for the UE with multiple MIM applications. Besides, longer t_R is effective in reducing the signaling load especially when the UTM mean interval is relatively low. This can guide network operators to set RRC release timer when reducing the signaling load.

4.2 UE energy consumption

The average energy consumption per unit time is given in Fig. 4. For the first two scenarios, when t_R is smaller than t_{KA}, the energy consumption decreases while t_R gets shorter or $1/\lambda$ gets longer as shown in Fig. 4(a, b). With the increase of t_{KA}, the energy consumption gets lower. But when t_{KA} increases to five times of $1/\lambda$, the energy consumption turns to be a constant as shown in Fig. 4(b). According to (13), the energy consumption depends on the average time for a UE to remain in RRC connected state per message $\overline{t_c}$ and the mean interval of joint messages $\overline{t_m}$. From Fig. 5(a), we can see that $\overline{t_c}$ doesn't change with t_{KA} when t_R is smaller than t_{KA}. It's $\overline{t_m}$ that changes with t_R as shown in Fig. 5(b). However, $\overline{t_m}$ increases with t_{KA} but will never surpass $1/\lambda$. Once t_{KA} surpasses five times of $1/\lambda$, $\overline{t_m}$ won't change any more, and the energy consumption will turn to be a constant accordingly. This can be easily understood according to (15). We can draw the conclusion that t_{KA} is upper bounded in the course of reducing UE energy consumption. The upper bound is five times of $1/\lambda$. This can be used for application operators to configure the KAM period. When t_R is larger than t_{KA}, the energy consumption is always 288 mAh, this can be explained by (14). For the UE with multiple MIM applications, longer t_R still brings higher energy consumption. Moreover, the energy consumption increases while $\min\{t_{KAi}\}$ or $\text{avg}\{t^n_{mi}\}$ gets shorter as shown in Fig. 4(c). Once t_R surpasses the $\min\{t_{KAi}\}$, energy consumption turns to 288mAh. The reason is the same with energy consumption in the case of the UE with single MIM application.

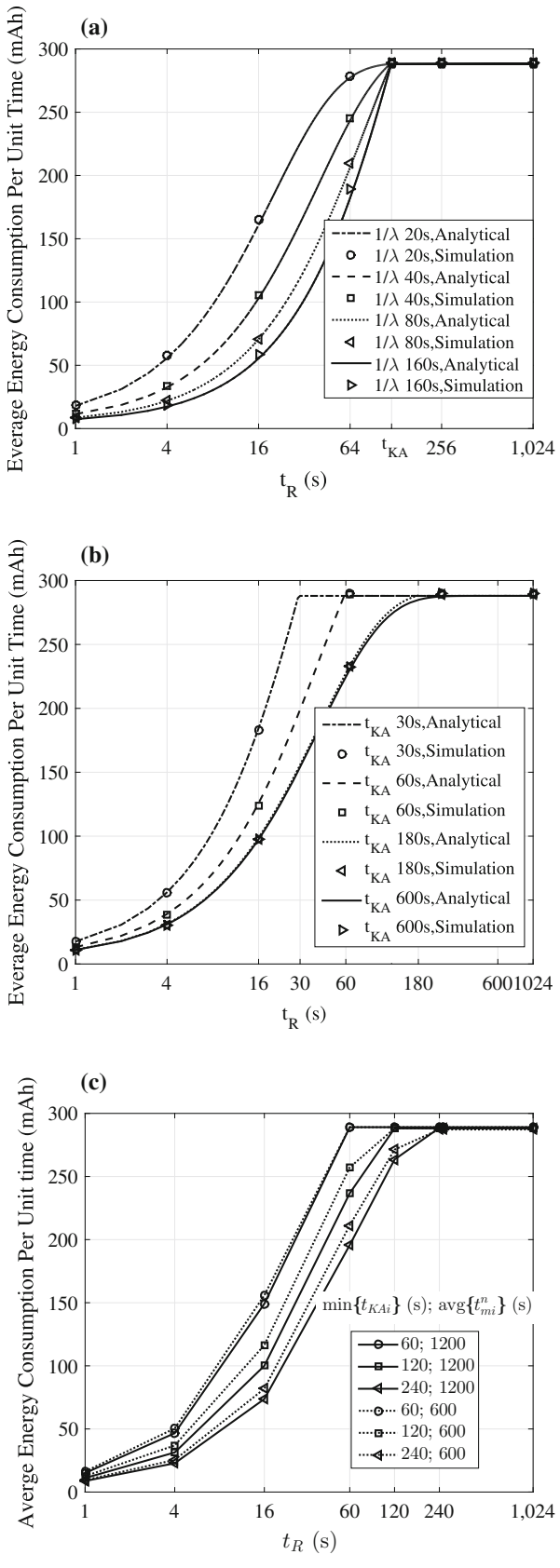


Fig. 4 Average UE energy consumption per message per unit time. **a** E (Scenario I). **b** E (Scenario II). **c** E (Scenario III)

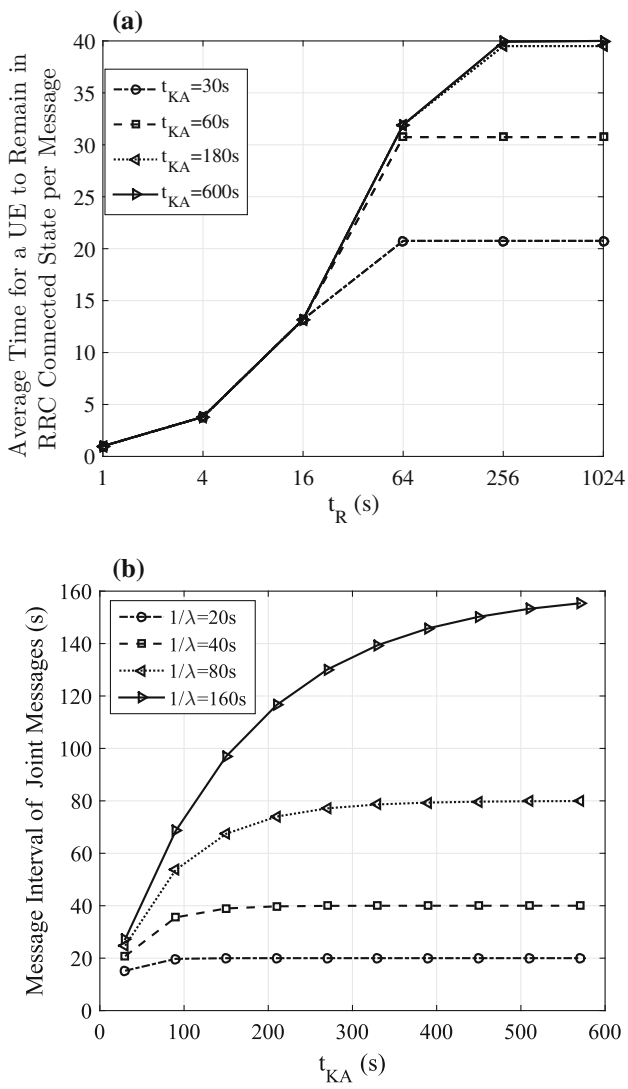


Fig. 5 Average time for a UE to remain in RRC connected state per message and mean interval of joint messages in scenario II. **a** $\overline{t_m}$ (Scenario II). **b** $\overline{t_m}$ (Scenario II)

5 Conclusions

In this paper, we have analyzed the impact of MIM applications on signaling load and UE energy consumption with different settings of RRC release timer, KAM period, and UTM mean interval. Analytical and simulation results show that the average signaling load per unit time decreases with increasing RRC release timer only when the timer is smaller than KAM period. Therefore, RRC release timer is upper bounded in the course of reducing signaling load. The upper bound is KAM period for the UE with single MIM application and the minimum KAM period for the UE with multiple MIM applications. Moreover, longer RRC release timer is particularly effective in decreasing the signaling load when the UTM mean interval of UE with

single application or the average UTM mean interval of UE with multiple applications is relatively low. In addition, the average UE energy consumption per message per unit time and the average signaling load per unit time decrease with increasing KAM period until the KAM period reaches five times of UTM mean interval. Therefore, KAM period can be optimized to reduce UE energy consumption and signaling load, but also is upper bounded. The upper bound is five times of UTM mean interval. These results can guide both network operators and MIM application developers to properly set control parameters for balancing the signaling load and UE energy consumption.

6 Future works

As mentioned in the previous section, signaling load S and UE energy consumption E have been formulated with the help of RRC release timer, KAM period, and UTM mean interval. We assume UTM mean interval is exponentially distributed, which is identified by 3GPP. However, some researchers argue that lognormal distribution fits UTM mean interval better. Our work can be further improved in deriving the general expressions of S and E regardless of the specific PDF of UTM message interval. Therefore, the impacts of MIM applications on signaling load and UE energy consumption can be presented more comprehensively.

Acknowledgments This work is sponsored by the National High Technology Research and Development Program of China (863 Program: 2015AA01A706); and by the Scientific Research Foundation of the Ministry of Education of China—China Mobile (No. MCM20150102).

Appendix 1: Derivation of the PDF of t_m when $0 \le t_m < t_{KA}$

t_m^n is a exponentially distributed random variable. Assume it is a sequence with infinite length of L. The number of times that a specific t_m^n appears in the whole sequence is $L \cdot (\lambda e^{-\lambda t_m^n})$. Therefore, the probability of the specific t_m^n when $t_m^n \ge 0$ can be rewrited as follows:

$$f_c(t_m^n) = \lambda e^{-\lambda t_m^n} = \frac{L \cdot (\lambda e^{-\lambda t_m^n})}{L}.$$

With different value of t_m^n, the respective probability can be obtained. Therefore, $f_c(t_m^n)$ can be used to denote the PDF of t_m^n. After the transformation of (1), the length of the new sequence, i.e., L' and the number of times that a specific t_m appears in the new sequence, i.e., $n(t_m)$ are no longer L and $L \cdot (\lambda e^{-\lambda t_m^n})$. But both of them can be expressed as the sum of piecewise functions of $f_c(t_m^n)$ when $0 \le t_m < t_{KA}$:

$$n(t_m) = \sum_{j=0}^{\left\lceil \frac{t_{max}}{t_{KA}} \right\rceil} L \cdot f_c(t_m + j \cdot t_{KA}),$$

$$L' = L + \sum_{j=1}^{\left\lceil \frac{t_{max}}{t_{KA}} \right\rceil} jL \cdot \left[\int_{jt_{KA}}^{(j+1)t_{KA}} f_c(t_m) dt_m - f_c(jt_{KA}) \right],$$

where t_{max} is the maximal value of t_m. Therefore, the PDF of t_m when $0 \le t_m < t_{KA}$ can be denoted as follows:

$$f(t_m) = \frac{n(t_m)}{l(t_m)} = \frac{\sum_{j=0}^{\left\lceil \frac{t_{max}}{t_{KA}} \right\rceil} L \cdot f_c(t_m + j \cdot t_{KA})}{L + \sum_{j=1}^{\left\lceil \frac{t_{max}}{t_{KA}} \right\rceil} jL \cdot \left[\int_{jt_{KA}}^{(j+1)t_{KA}} f_c(t_m) dt_m - f_c(jt_{KA}) \right]}.$$

Since t_{max} is much longer than t_{KA}, we have $t_{max} = N \cdot t_{KA}$. Thus, $f(t_m)$ can be simplified as follows:

$$f(t_m) = \frac{\sum_{j=0}^{N} f_c(t_m + j \cdot t_{KA})}{1 + \sum_{j=1}^{N} j \left[\int_{jt_{KA}}^{(j+1)t_{KA}} f_c(t_m) dt_m - f_c(jt_{KA}) \right]}$$

$$= \frac{\frac{\lambda \cdot e^{-\lambda \cdot t_m}}{1 - e^{-\lambda \cdot t_{KA}}}}{\frac{1}{1 - e^{-\lambda \cdot t_{KA}}} - Ne^{-\lambda \cdot (N+1) \cdot t_{KA}} - \left[\frac{\lambda \cdot e^{-\lambda \cdot t_{KA}}}{(1 - e^{-\lambda \cdot t_{KA}})^2} - \frac{N\lambda \cdot e^{-\lambda \cdot (N+1) \cdot t_{KA}}}{1 - e^{-\lambda \cdot t_{KA}}} \right]}.$$

Generally, N is large enough, $f(t_m)$ can be further simplified:

$$f(t_m) = \frac{\frac{\lambda \cdot e^{-\lambda \cdot t_m}}{1 - e^{-\lambda \cdot t_{KA}}}}{\frac{1}{1 - e^{-\lambda \cdot t_{KA}}} - \frac{\lambda \cdot e^{-\lambda \cdot t_{KA}}}{(1 - e^{-\lambda \cdot t_{KA}})^2}} = \frac{\lambda \cdot e^{-\lambda \cdot t_m}}{1 - \frac{\lambda \cdot e^{-\lambda \cdot t_{KA}}}{1 - e^{-\lambda \cdot t_{KA}}}}.$$

Since t_{KA} is much bigger than $1/\lambda$ in most case, $\frac{\lambda \cdot e^{-\lambda \cdot t_{KA}}}{1 - e^{-\lambda \cdot t_{KA}}}$ is negligibly small. Therefore, the final expression of $f(t_m)$ when $0 \le t_m < t_{KA}$ is:

$$f(t_m) = \lambda e^{-\lambda t_m},$$

which is the same with $f_c(t_m^n)$.

Appendix 2: Calculation of the mean value of t_m when $0 \le t_m \le x (0 \le x < t_{KA})$

When $0 \le t_m < t_{KA}$, the PDF of t_m is typical exponentially distributed. In addition of $0 \le x < t_{KA}$, the PDF of t_m when $0 \le t_m \le x$ is head-truncated exponentially distributed, which is denoted by $f_t(t_m)$. Since the shape of $f_t(t_m)$ is the same with $f(t_m)$, $f_t(t_m)$ can be expressed as follows:

$$f_t(t_m) = \alpha f(t_m),$$

where α can be calculated by:

$$\int_0^x \alpha \cdot \lambda e^{-\lambda t_m} dt_m = 1,$$

Therefore, α can be obtained accordingly:

$$\alpha = \frac{1}{1 - e^{-\lambda x}},$$

Then, the mean value of t_m when $0 \le t_m \le x (0 \le x < t_{KA})$, i.e., $h(x)$ can be calculated as follows:

$$h(x) = \int_0^x t_m \cdot f_t(t_m) dt_m = \frac{\left(x + \frac{1}{\lambda}\right) \cdot e^{-\lambda x} - \frac{1}{\lambda}}{e^{-\lambda x} - 1}.$$

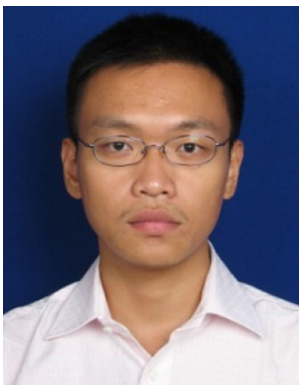
References

1. Cuadrado, F., & Dueñas, J. C. (2012). Mobile application stores: Success factors, existing approaches, and future developments. *IEEE Communication Magazine*, *50*(11), 160–167.
2. Karimiyazdi, R., & Mokhber, M. (2015). Improving viral marketing campaign via mobile instant messaging (MIM) applications. In *Proceedings of 1st world virtual conference social and behavioural sciences* (pp. 1–13).
3. Deloitte. (2015). Short message services versus instant messaging: Value versus volume. http://www2.deloitte.com/tw/tc/pages/technology-media-and-telecommunications/articles/2014prediction-short-messaging-svcs-vs-instant-messaging.html. Accessed 03 June 2015.
4. Zhou, X., Zhao, Z., Li, R., Zhou, Y., Palicot, J., & Zhang, H. (2014). Understanding the nature of social mobile instant messaging in cellular networks. *IEEE Communications Letters*, *18*(3), 389–392.
5. Li, R., Zhao, Z., Qi, C., Zhou, X., Zhou, Y., & Zhang, H. (2015). Understanding the traffic nature of mobile instantaneous messaging in cellular networks: A revisiting to α-stable models. *IEEE Access*, *3*, 1416–1422.
6. Yang, C. (2011). Weather the signaling storm. *Huawei Communicate*, *61*, 18–20.
7. Choi, Y., Yoon, C. H., Kim, Y. S., Heo, S. W., & Silvester, J. A. (2014). The impact of application signaling traffic on public land mobile networks. *IEEE Communications Magazine*, *52*(1), 166–172.
8. Abdelrahman, O., & Gelenbe, E. (2014). Signalling storms in 3G mobile networks. In *2014 IEEE international conference on communications (ICC)* (pp. 1017–1022). IEEE.
9. Aucinas, A., Vallina-Rodriguez, N., Grunenberger, Y., Erramilli, V., Papagiannaki, K., Crowcroft, J., et al. (2013). Staying online while mobile: The hidden costs. In *Proceedings of the ninth ACM conference on emerging networking experiments and technologies* (pp. 315–320). ACM.
10. Gupta, M., Jha, S., Koc, A., & Vannithamby, R. (2013). Energy impact of emerging mobile Internet applications on LTE networks: Issues and solutions. *IEEE Communication Magazine*, *51*(2), 90–97.
11. Ou, Z., Dong, S., Dong, J., Nurminen, J. K., Ylä-Jääski, A., & Wang, R. (2013). Characterize energy impact of concurrent network-intensive applications on mobile platforms. In *Proceedings of the eighth ACM international workshop on Mobility in the evolving internet architecture* (pp. 23–28). ACM.
12. Balasubramanian, N., Balasubramanian, A., & Venkataramani, A. (2009). Energy consumption in mobile phones: A measurement study and implications for network applications. In *Proceedings of the 9th ACM SIGCOMM conference on Internet measurement conference* (pp. 280–293). ACM.
13. Zhou, Z., Dong, M., Ota, K., Wu, J., & Sato, T. (2014). Energy efficiency and spectral efficiency tradeoff in device-to-device (D2D) communications. *IEEE Wireless Communications Letters*, *3*(5), 485–488.
14. Xu, Y., Hu, R. Q., Qian, Y., & Znati, T. (2016). Video quality-based spectral and energy efficient mobile association in heterogeneous wireless networks. *IEEE Transactions on Communications*, *64*(2), 805–817.
15. Schwartz, C., Hoßfeld, T., Lehrieder, F., & Tran-Gia, P. (2013). Angry apps: The impact of network timer selection on power consumption, signalling load, and web qoe. *Journal of Computer Networks and Communications*, *2013*, 1–13.
16. Zhang, Z., Zhao, Z., Guan, H., Miao, D., & Tan, Z. (2013). Study of signaling overhead caused by keep-alive messages in LTE network. In *2013 IEEE 78th vehicular technology conference (VTC Fall)* (pp. 1–5). IEEE.
17. Nokia Corporation, Nokia Siemens Networks. (2011). *Power consumption and signalling load for background traffic*. 3GPP TSG-RAN WG2 LTE contribution R2-115931.
18. Jia, Y., Zhang, Y., Liang, L., & Zhou, S. (2015). An energy-efficient system signaling control method based on mobile application traffic. In *2015 IEEE international conference on communications (ICC)* (pp. 232–237). IEEE.
19. Puttonen, J., Virtej, E., Keskitalo, I., & Malkamäki, E. (2012). On LTE performance trade-off between connected and idle states with always-on type applications. In *2012 IEEE 23rd international symposium on personal indoor and mobile radio communications (PIMRC)* (pp. 981–985). IEEE.
20. Nokia Corporation, Nokia Siemens Networks. (2012). *Further results on network signalling load and UE power consumption*. 3GPP TSG-RAN WG2 LTE contribution R2-120367.
21. Foddis, G., Garroppo, R. G., Giordano, S., Procissi, G., Roma, S., & Topazzi, S. (2015). LTE traffic analysis for signalling load and energy consumption trade-off in mobile networks. In *2015 IEEE international conference on communications (ICC)* (pp. 6005–6010). IEEE.
22. The 3rd Generation Partnership Project (3GPP); Technical Specification Group GSM/EDGE; Radio Access Network (2014). GERAN study on mobile data applications. TR 43.802, v12.0.0.
23. The 3rd Generation Partnership Project (3GPP); Technical Specification Group Radio Access Network; Evolved Universal Terrestrial Radio Access (E-UTRA) (2016). Radio Resource Control (RRC) Protocol specification. TS 36.331, v13.1.0.
24. The 3rd Generation Partnership Project (3GPP); Technical Specification Group Radio Access Network; LTE Radio Access Network (RAN) (2012). Enhancements for Diverse Data Applications. TR 36.822, v11.0.0.

Yunjian Jia received his B.S. degree from Nankai University, China, and his M.E. and Ph.D. degrees in Engineering from Osaka University, Japan, in 1999, 2003 and 2006, respectively. From 2006 to 2012, he was with Central Research Laboratory, Hitachi, Ltd., where he engaged in research and development on wireless networks, and also contributed to LTE/LTE-Advanced standardization in 3GPP. He is now a Professor at the College of Communication Engineering, Chongqing University, Chongqing, China. He is the author of more than 60 published papers, and 30 granted patents. His research interests include future radio access technologies, mobile networks, and IoT. He has won several prizes from industry and academia including the IEEE Vehicular Technology Society Young Researcher Encouragement Award, the IEICE Paper Award, the Yokosuka Research Park R&D Committee YRP Award, and the Top 50 Young Inventors of Hitachi. Moreover, he was a research fellowship award recipient of International Communication Foundation in 2004, and Telecommunications Advancement Foundation Japan in 2005.

Yu Zhang received her B.E. degree in Communication Engineering from Chongqing University, China, in 2014. She is currently working toward her M.E. degree in Information and Communication Engineering in the same university. Her research interests include radio resource management and optimization for mobile networks, with emphasis on quality-of-experience provision.

Liang Liang received her B.E. and M.E. degrees from the Southwest University of Science and Technology (SWUST), China, in 2003 and 2006, respectively, and the Ph.D. degree in communication and information system from the University of Electronic Science and Technology of China (UESTC) in 2012. She is currently a Lecturer in College of Communication Engineering, Chongqing University, Chongqing, China. Her research interests include wireless communication and optimization, green radio, and wireless sensor networks.



Weiyang Xu received his B.S.E. degree and M.S.E. degree from the Xi'an Jiao Tong University, Xi'an, China, in 2004 and 2007, respectively, and the Ph.D. degree from Fudan University, Shanghai, China, in 2010. He is currently an Associate Professor in the College of Communications Engineering at Chongqing University, Chongqing, China. In 2014, he was a visiting scholar at University of Southern Queensland, Australia. His research interests include channel coding, OFDM, massive MIMO and cognitive radio techniques.

Sheng Zhou received the B.E. and Ph.D. degrees in electronic engineering from Tsinghua University, Beijing, China, in 2005 and 2011, respectively. From January to June 2010, he was a visiting student at the Wireless System Lab, Department of Electrical Engineering, Stanford University, Stanford, CA, USA. He is currently an Assistant Professor with the Department of Electronic Engineering, Tsinghua University. His research interests include cross-layer design for multiple antenna systems, cooperative transmission in cellular systems, and green wireless communications. He coreceived the Best Paper Award at the Asia-Pacific Conference on Communication in 2009 and 2013, the 23th IEEE International Conference on Communication Technology in 2011, and the 25th International Tele-traffic Congress in 2013.