Energy-Delay Tradeoff Analysis of User State Transition Mechanism for Mobile Web Services

Hyun-Ho Choi[a], Ki-Ho Lee[b], Jung-Ryun Lee[c,*]

[a]*Department of Electrical, Electronic and Control Engineering, Institute for Information Technology Convergence, Hankyong National University, Republic of Korea*
[b]*Mobile R&D Laboratory, KT, Republic of Korea*
[c]*School of the Electrical Engineering, Chung-Ang University, Republic of Korea*

Abstract

From the perspective of mobile users, the most considerable experiences are service delay and battery lifetime, both of which are closely related to the user state transition mechanism adopted in most mobile wireless systems. In this paper, the user state transition mechanism with more than simple two states (active/idle or on/off states) is modeled through Markov chain. Then, its performance is numerically analyzed with respect to the energy consumption and the activation delay by considering the bursty traffic attributes of mobile web services. As a result, we derive the energy-delay tradeoff curves, which show an achievable performance bound of the user state transition mechanism. The derived energy-delay tradeoff curve provides a guideline for network operators to suitably apply the user state transition mechanism so as to minimize the energy consumption of a mobile node while guaranteeing the delay requirement of the service.

*Corresponding author
Email addresses: hhchoi@hknu.ac.kr (Hyun-Ho Choi), kiholee@kt.com (Ki-Ho Lee), jrlee@cau.ac.kr (Jung-Ryun Lee)

Preprint submitted to Journal of Network and Computer Applications September 24, 2013

Keywords: energy-delay tradeoff, user state transition, state management, power management, mobile web services.

1. Introduction

The use of smartphones is becoming popular and their market share is rapidly increasing. Smartphones are used for various multimedia applications that require high data rates, but major applications involve the transmission of small-sized and frequently updated data, such as in web browsing, email synchronization, stock portfolio updates, social media, etc. (GSM Association, 2010). Such mobile services have common traffic attributes, as shown in Fig. 1 (Pries et al., 2012). First, a small-sized request packet is transmitted via the uplink to request the download of content. As a response, bursty data packets containing information are delivered via the downlink. Then, after a certain time period during which the user reads the downloaded content, the next request packet is generated again. These basic request and response packet transmissions for a single page download take place continuously by human interactions until the session ends.

From the viewpoint of the quality of experience (QoE) of mobile users, the most important considerations are service delay and battery lifetime, because users are very sensitive to how fast a new page loads and how long the battery will last without recharging. These energy and delay performances are closely related to the user state transition mechanism in mobile wireless systems. Therefore, to provide the user a better QoE in terms of energy and delay, it is important to analyze the behavior of the user state transition mechanism and to provide network operators with the method to appropri-

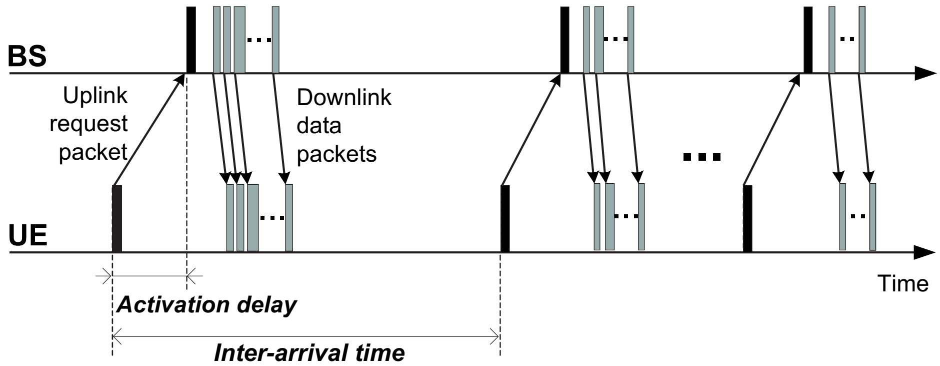


Figure 1: Traffic attributes of mobile web services.

ately apply the user state transition mechanism by considering the attribute of various applications. Typically, the 3G cellular system defines the user state transition mechanism, which manages the state of the user equipment (UE) in order to use radio resources and battery power efficiently (3GPP TS 25.331, 2012). Its basic operation is that when user data is generated, a dedicated or common radio channel is allocated to the UE and data is transmitted promptly. When there is no traffic arrival during some pre-determined time duration, the radio unit of the UE powers off and its energy consumption is minimized. In the power-off state, energy use is at a minimum, but the *activation delay*[1] is increased because the transmission of the uplink request packet requires heavy signaling procedures for random access and connection setup. On the other hand, in the power-on state, such signaling procedures are omitted so that the activation delay is shortened considerably.

[1]In mobile web services, the activation delay is a dominant delay factor varying with the user state and is defined as the time interval from the time the uplink request packet occurs at the UE to its arrival time at the base station (BS), as shown in Fig. 1.

It has been recognized that a tradeoff exists between the energy conservation of a UE and the delay performance of the transmitted packets. In many previous literatures, the tradeoff between the energy consumption and the downlink packet buffering delay has been investigated (Lee et al., 2007; Sarkar et al., 2004; Nga et al., 2007; Vuyst et al., 2009). This tradeoff arises from the fact that the longer the UE stays in the idle state, the more power is saved, but an additional buffering delay happens. A probabilistic sleep interval decision algorithm has been previously proposed to guarantee the delay requirement according to the energy-delay tradeoff characteristics (Lee et al., 2007). The sleep and awake states in the downlink WLAN environment have also been considered, and the timing and duration of sleep states were optimized to minimize the buffering delay (Sarkar et al., 2004). Another studies analyzed the IEEE 802.16e sleep-mode mechanism from the perspective of energy-delay tradeoff and enhanced the standard power-saving scheme (Nga et al., 2007; Vuyst et al., 2009). The tradeoff between the power consumption and the mouth-to-ear (end-to-end) delay has also been studied (Choi et al., 2009; Jang et al., 2010). Therein, the sleep threshold and interval were controlled in order to minimize the power consumption while satisfying the delay constraint. A tradeoff issue between the energy consumption and wake-up delay has been mentioned as well (Chiasserini et al., 2003). It was observed that nodes in deeper sleep states consume less energy while asleep, but incur a longer delay and higher energy cost to awaken. The energy-delay tradeoff when the smartphone uses multiple wireless interfaces, viz. 3G, EDGE and WiFi, has been addressed (Ra et al., 2010). Therein, the tradeoff between energy consumption and connection delay occurs because the UE delays the data transfers until a lower-energy WiFi connection becomes available. The possibility for tradeoff between video quality and power saving in the receiver has been demonstrated (Ukhanova et al., 2012). This work took the power-delay tradeoff into account and showed how it influences the quality of video transmission, such as data bit rates and peak signal-to-noise ratio (PSNR). Furthermore, some numerical results have been presented that are related to the user state transition mechanism of the 3GPP standard (Chung et al., 2002; Yeh et al., 2004). However, these works did not provide any analytical interpretation from the perspective of the energy-delay tradeoff.

While previous works have investigated the energy-delay tradeoff considering delay metrics, such as downlink buffering delay, mouth-to-ear delay, wake-up delay, and connection delay, our work analyzes another relationship between energy consumption of mobile users and uplink activation delay, and thereby focuses on the QoE (i.e., service delay and battery lifetime) of mobile users. In addition, the previous analysis mainly considered simple two-state (active/idle or on/off) transition mechanisms, but our analysis considers all user states defined in the standard, such as DCH, PCH, FACH, and IDLE states (as shown in Fig. 2), which have so far not been addressed to the best of our knowledge. In this paper, we analyze the user state transition mechanism of the 3G system with respect to the energy-delay tradeoff by considering the bursty traffic attributes of mobile web services. Our main contribution is to characterize the performances of energy consumption and activation delay as a simple energy-delay tradeoff curve. This tradeoff curve reveals the fundamental limits and characteristics of the user state transition mechanism according to operational parameters; thus, we present an optimal operation strategy for efficient state management according to QoE requirements.

The rest of this paper is organized as follows. In Section 2, the standard user state transition mechanism is explained in detail. In Sections 3, the state transition mechanism is modeled and its performances are evaluated with respect to the average energy consumption and activation delay. In Section 4, analysis and simulation results are presented. Section 5 concludes this paper.

2. User State Transition Mechanism

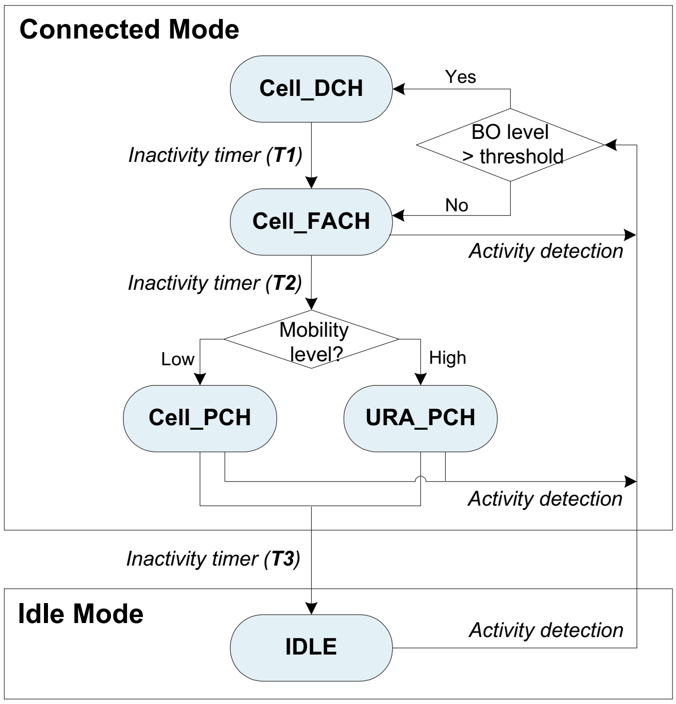


Figure 2: Mechanism of user state transition.

Fig. 2 illustrates the considered user state transition mechanism (3GPP

TS 25.331, 2012). It is mainly divided into two modes, idle and connected, and the corresponding state(s) exist for each mode. When the UE is powered on, it enters the IDLE state in the idle mode. In the IDLE state, the UE is attached to the network but is not actively engaged in data transfer. If there is a packet to transmit, the UE establishes a radio connection and switches to connected mode, which consists of the following four service states: Cell_DCH (Dedicated channel), Cell_FACH (Fast Access channel), Cell_PCH (Paging channel), and URA_PCH (UTRAN Registration Area Paging channel). In the Cell_DCH state, user data is transferred through a DCH. In the Cell_FACH state, small data is carried through common channels, such as FACH for a downlink transmission and RACH (Random Access channel) for an uplink transmission. In the Cell_PCH or URA_PCH state, the UE does not transfer data but discontinuously listens to the PCH in order to check whether or not any pending data is present in the BS. The difference between the two PCH states is only in the criterion of the location update. A UE in the URA_PCH state updates its location information when the URA (defined as a set of cells) is changed, whereas a UE in the Cell_PCH state performs a location update procedure when the serving cell is changed. Therefore, the URA_PCH state is more suitable than the Cell_PCH state for high mobility users.[2]

The transition between any two states depends on the *buffer occupancy (BO) level* and the *inactivity timers*. The transition to the lower states occurs

[2]User mobility is out of the scope of this paper, so we do not distinguish between the terms Cell and URA in this paper. Hereafter, we simply denote the state name without the Cell or URA prefix.

when the BO level is zero and the inactivity timer expires. The inactivity timers, T1, T2 and T3 control the transitions from DCH to FACH, from FACH to PCH, and from PCH to IDLE, respectively. Conversely, from the FACH, PCH or IDLE state, the transition back to the DCH or FACH state is carried out whenever user activity (i.e., the transmission of a request packet) is detected. The selection between the DCH and FACH states depends on the amount of data to be transmitted; the DCH state is selected when the BO level exceeds a predetermined threshold value. Otherwise, the FACH state is selected.

The state transition mechanism has a direct impact on the end-user experience. The battery lifetime is longest when the device remains in the IDLE state. The power consumption of UE in the FACH state is roughly 40 percent of that in the DCH state, and the PCH and IDLE states use about 1$\sim$2 percent of the DCH state power consumption (GSM Association, 2010). On the other hand, in the IDLE state the UE experiences the longest activation delay, which is made up of the radio wake-up time, the connection setup time, and the request packet transmission time. However, when the UE stays in the PCH state, the signaling procedures for the connection setup are not needed, so the activation delay is significantly decreased. When the UE is in the FACH or DCH state, the activation delay depends only on the request packet transmission time. The UE in FACH state involves random access delay for the uplink packet transmission, while the UE in DCH state transmits uplink packets directly through a DCH without delay (Holma et al., 2004). Therefore, the fundamental tradeoff is summarized as follows: the longer the UE remains in the upper states, the more power is consumed;

however, the user experiences a shorter delay.

3. Performance Analysis

As explained, state transition is triggered either by timer expiration or packet arrival. In general, the inter-arrival time of request packets generated in non real-time applications is assumed to be exponentially distributed with mean $1/\lambda$ (IEEE 802.16m, 2009; Pries et al., 2012), so the state transition mechanism can be modeled as a discrete-time Markov chain (DTMC), as shown in Fig. 3. There are four states and $P_{A,B}$ denotes the state transition probability from state A to state B. All transitions from a state to its following right-hand state happen by the expiration of inactivity timers, while the reverse transitions happen by the arrivals of request packet.

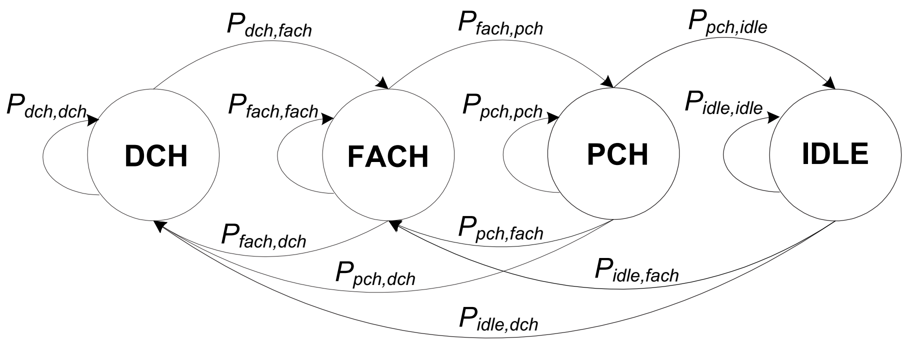


Figure 3: Discrete-time Markov chain model of state transition mechanism.

First of all, we try to obtain the probabilities of state transitions triggered by the timer expirations ($P_{dch,fach}$, $P_{fach,pch}$ and $P_{pch,idle}$). Since these transitions follow the same mechanism, we may consider only two states without loss of generality. Suppose there are arbitrary two states, A and B. Let T_u be the unit time of the state transition mechanism, and let T be the value of the inactivity timer for the transition from state A to state B. The timer value T is expressed as $T = nT_u$ for a given integer n. Thus, when no uplink request packet is generated at the UE in state A during n consecutive time slots, the UE goes into state B. Based on this slotted operation, we subdivide state A into n substates, which is denoted by A(i) for $i = 0, 1, 2, \cdots, n-1$. That is, a UE entering state A goes into substate A(i) if it does not generate any request packets during i successive time slots. Therefore, the state transition mechanism between states A and B are represented as an $(n+1)$-substate model of {A(i), B | $i = 0, 1, 2, \cdots, n-1$}, which is drawn in Fig. 4. The probability that state A(i) moves forward to state A(i+1) is obtained by

$$P_\beta = Pr\{\text{no packet arrival during } T_u\} = 1 - \int_0^{T_u} f(x)dx = e^{-\lambda T_u} \tag{1}$$

where $f(x) = \lambda e^{-\lambda x}$ is the probability density function of the packet inter-arrival time. On the other hand, if there is a packet arrival during T_u in any state, the state moves back to state A(0), the probability of which is given by

$$P_\alpha = Pr\{\text{packet arrival during } T_u\} = 1 - P_\beta = 1 - e^{-\lambda T_u}. \tag{2}$$

Consequently, the state transition matrix $\mathbf{M}$ with a size of $(n+1) \times (n+1)$ is described as

$$\mathbf{M} = \begin{bmatrix} 1-e^{-\lambda T_u} & e^{-\lambda T_u} & 0 & \cdots & 0 \\ 1-e^{-\lambda T_u} & 0 & e^{-\lambda T_u} & & 0 \\ 1-e^{-\lambda T_u} & 0 & 0 & \ddots & \vdots \\ \vdots & \vdots & \vdots & & e^{-\lambda T_u} \\ 1-e^{-\lambda T_u} & 0 & 0 & \cdots & e^{-\lambda T_u} \end{bmatrix}. \tag{3}$$

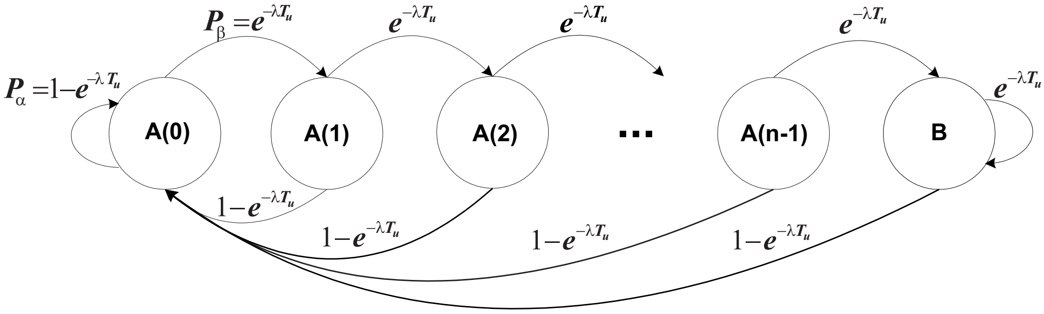


Figure 4: DTMC model of state transitions between only two states, A and B.

Let π_i be the steady state probability and $\Pi = [\pi_0 \; \pi_1 \; \cdots \; \pi_n]$ be the probability vector. By the balance equation $\Pi = \Pi \cdot \mathbf{M}$ and the normalized condition $\sum_{i=0}^{n} \pi_i = 1$, π_i is calculated as

$$\pi_i = \begin{cases} (1 - e^{-\lambda T_u}) e^{-\lambda i T_u} & \text{for } i = 0, 1, 2, \cdots, n-1 \\ e^{-\lambda n T_u} & \text{for } i = n. \end{cases} \tag{4}$$

Now, in order to obtain the transition probability from state A to state B, we define some parameters as follows:

- N: The total number of time slots during the whole session time.
- N_A: The number of time slots during which a UE stays in state A.
- $N_{A,B}$: The number of time slots at which a UE changes state A into state B.

According to the above definition, we develop the following formulas:

$$\begin{aligned}
N_{A(i)} &= N \cdot \pi_i = N(1-e^{-\lambda T_u})e^{-\lambda i T_u} \quad \text{for} \quad i = 0, 1, 2, \cdots, n-1 && (5)\\
N_A &= \sum_{i=0}^{n-1} N_{A(i)} = N(1-e^{-\lambda T_u}) \sum_{i=0}^{n-1} e^{-\lambda i T_u} = N(1-e^{-\lambda n T_u}) && (6)\\
N_{A,B} &= N_{A(n-1),B} = N_{A(n-1)} P_\beta \\
&= N(1-e^{-\lambda T_u})e^{-\lambda (n-1) T_u} e^{-\lambda T_u} = N(1-e^{-\lambda T_u})e^{-\lambda n T_u}. && (7)
\end{aligned}$$

Using (6) and (7), the transition probability from state A to state B is obtained by

$$P_{A,B} = \frac{N_{A,B}}{N_A} = \frac{N(1-e^{-\lambda T_u})e^{-\lambda n T_u}}{N(1-e^{-\lambda n T_u})} = \frac{(1-e^{-\lambda T_u})e^{-\lambda T}}{1-e^{-\lambda T}}. \qquad (8)$$

As mentioned, because the transition mechanisms by three timer expirations are identical, three state transition probabilities, $P_{dch,fach}$, $P_{fach,pch}$ and $P_{pch,idle}$, can be derived by changing the inactivity timer value T from (8).

On the other hand, the probabilities of state transitions triggered by the request packet arrival are just given by the packet arrival probability, $P_\alpha = 1 - e^{-\lambda T_u}$ from (2), but we should additionally consider the amount of downlink data packets generated because the choice of state transition to DCH or FACH is dependent on the volume of downlink traffic. To do this, we introduce a parameter w which denotes the probability of the amount of data packets generated being greater than the predetermined buffer threshold. So, when a new request packet occurs in any state, the state goes into the DCH state with the probability w and the FACH state with the probability $(1-w)$. Note that in practice, the parameter w is controlled by the buffer threshold. That is, w increases as the buffer threshold decreases. From this concept and equation (8), the state transition probabilities in each state are summarized as follows:

- In DCH state:

$$P_{dch,fach} = \frac{(1-e^{-\lambda T_u})e^{-\lambda T_1}}{1-e^{-\lambda T_1}} \quad (9)$$

$$P_{dch,dch} = 1 - P_{dch,fach} \quad (10)$$

- In FACH state:

$$P_{fach,pch} = \frac{(1-e^{-\lambda T_u})e^{-\lambda T_2}}{1-e^{-\lambda T_2}} \quad (11)$$

$$P_{fach,dch} = w(1-e^{-\lambda T_u}) \quad (12)$$

$$P_{fach,fach} = 1 - P_{fach,dch} - P_{fach,pch} \quad (13)$$

- In PCH state:

$$P_{pch,idle} = \frac{(1-e^{-\lambda T_u})e^{-\lambda T_3}}{1-e^{-\lambda T_3}} \quad (14)$$

$$P_{pch,dch} = w(1-e^{-\lambda T_u}) \quad (15)$$

$$P_{pch,fach} = (1-w)(1-e^{-\lambda T_u}) \quad (16)$$

$$P_{pch,pch} = 1 - P_{pch,idle} - P_{pch,dch} - P_{pch,fach} \quad (17)$$

- In IDLE state:

$$P_{idle,dch} = w(1-e^{-\lambda T_u}) \quad (18)$$

$$P_{idle,fach} = (1-w)(1-e^{-\lambda T_u}) \quad (19)$$

$$P_{idle,idle} = 1 - P_{idle,dch} - P_{idle,fach} \quad (20)$$

where T_1, T_2 and T_3 are the inactivity timer values for the state transitions from DCH to FACH, from FACH to PCH, and from PCH to IDLE, respectively. So, the state transition matrix of Fig. 3 is expressed as

$$\mathbf{M}' = \begin{bmatrix} P_{dch,dch} & P_{dch,fach} & 0 & 0 \\ P_{fach,dch} & P_{fach,fach} & P_{fach,pch} & 0 \\ P_{pch,dch} & P_{pch,fach} & P_{pch,pch} & P_{pch,idle} \\ P_{idle,dch} & P_{idle,fach} & 0 & P_{idle,idle} \end{bmatrix}. \tag{21}$$

By solving the balance equation, the steady state probability is calculated as

$$\pi_{dch} = \frac{w(1-e^{-\lambda T_1})}{w+(1-w)e^{-\lambda T_1}} \tag{22}$$

$$\pi_{fach} = \frac{e^{-\lambda T_1}(1-e^{-\lambda T_2})}{w+(1-w)e^{-\lambda T_1}} \tag{23}$$

$$\pi_{pch} = \frac{e^{-\lambda(T_1+T_2)}(1-e^{-\lambda T_3})}{w+(1-w)e^{-\lambda T_1}} \tag{24}$$

$$\pi_{idle} = \frac{e^{-\lambda(T_1+T_2+T_3)}}{w+(1-w)e^{-\lambda T_1}}. \tag{25}$$

Let $\mathcal{E}_i$ be the energy consumption when the UE stays in state i for $i \in \{dch, fach, pch, idle\}$. Considering the time slots during which the UE stays in state i (N_i) among the total time slots of the session (N), the average energy consumption of UE per unit time is obtained by

$$\begin{aligned} \overline{\mathcal{E}} &= \frac{N_{dch}\mathcal{E}_{dch} + N_{fach}\mathcal{E}_{fach} + N_{pch}\mathcal{E}_{pch} + N_{idle}\mathcal{E}_{idle}}{N} \\ &= \frac{N\pi_{dch}\mathcal{E}_{dch} + N\pi_{fach}\mathcal{E}_{fach} + N\pi_{pch}\mathcal{E}_{pch} + N\pi_{idle}\mathcal{E}_{idle}}{N} \\ &= \pi_{dch}\mathcal{E}_{dch} + \pi_{fach}\mathcal{E}_{fach} + \pi_{pch}\mathcal{E}_{pch} + \pi_{idle}\mathcal{E}_{idle}. \end{aligned} \tag{26}$$

Furthermore, let $\mathcal{D}_i$ be the cost of activation delay required when the UE stays in state i for $i \in \{dch, fach, pch, idle\}$. The average activation delay depends on which state the UE remains in at the moment that the request packet is generated. During the session period (i.e., N time slots), the total number of request packet arrivals is NP_α. Accordingly, when the UE stays in the DCH or FACH state, the number of request packet arrivals becomes $N_{dch}P_\alpha$ and $N_{dch}P_\alpha$, respectively. On the other hand, when the UE is in the PCH or IDLE state, the request packet arrival induces the immediate state transition into the DCH or FACH state, and so the number of request packet arrivals for the UE in the PCH or IDLE state is given by $(N_{pch,dch} + N_{pch,fach})$ and $(N_{idle,dch} + N_{idle,fach})$, respectively. By using $N_{A,B}=N_A P_{A,B}$ from (8), $(N_{pch,dch} + N_{pch,fach})=N_{pch}(P_{pch,dch} + P_{pch,fach})$ and $(N_{idle,dch} + N_{idle,fach})=N_{idle}(P_{idle,dch} + P_{idle,fach})$ are derived. Therefore, the average activation delay is calculated as

$$\begin{aligned}
\overline{\mathcal{D}} &= \frac{N_{dch}P_\alpha\mathcal{D}_{dch} + N_{fach}P_\alpha\mathcal{D}_{fach}}{NP_\alpha} \\
&\quad + \frac{N_{pch}(P_{pch,dch} + P_{pch,fach})\mathcal{D}_{pch} + N_{idle}(P_{idle,dch} + P_{idle,fach})\mathcal{D}_{idle}}{NP_\alpha} \\
&= \frac{N_{dch}P_\alpha\mathcal{D}_{dch} + N_{fach}P_\alpha\mathcal{D}_{fach} + N_{pch}P_\alpha\mathcal{D}_{pch} + N_{idle}P_\alpha\mathcal{D}_{idle}}{NP_\alpha} \\
&= \frac{N\pi_{dch}P_\alpha\mathcal{D}_{dch} + N\pi_{fach}P_\alpha\mathcal{D}_{fach} + N\pi_{pch}P_\alpha\mathcal{D}_{pch} + N\pi_{idle}P_\alpha\mathcal{D}_{idle}}{NP_\alpha} \\
&= \pi_{dch}\mathcal{D}_{dch} + \pi_{fach}\mathcal{D}_{fach} + \pi_{pch}\mathcal{D}_{pch} + \pi_{idle}\mathcal{D}_{idle}. \qquad (27)
\end{aligned}$$

4. Results and Discussions

Table 1 shows the costs of energy consumption and activation delay in each state used in our work. We employ relative energy costs based on the fact that the PCH, FACH and DCH states consume energy approximately 2, 40 and 100 times more than the IDLE state, respectively (GSM Association, 2010). For the costs of activation delay, we use statistical delay data measured from the KT (Korea Telecom) UMTS networks, which also match well with previously reported measurement results (Perala et al., 2009). T_u

is set to 100 ms according to the 3GPP standard (3GPP TS 25.331, 2012) and the average packet inter-arrival time ($1/\lambda$) is varied within the range of 1$\sim$30 s (IEEE 802.16m, 2009). The default value of parameter w is set to 0.5, which means that the probability of state transition into DCH or FACH as a response to the request packet arrival is divided evenly. For evaluation, we suppose six feasible state management scenarios, as shown in Table 2. Notice that the states used in each scenario are varied by setting each inactivity timer to zero or infinity. That is, $T_1 = 0$, $T_2 = 0$ and $T_3 = 0$ do not involve the DCH, FACH and PCH states, respectively, and $T_2 = \infty$ and $T_3 = \infty$ do not involve the PCH and IDLE states, respectively. Except six scenarios in Table 2, other scenarios are not considered due to their impracticality. We perform Monte-Carlo experiments using MATLAB to validate the numerical analysis. By using MATLAB, we implement the standard operational procedures of the user state transition mechanism explained in Section 2. In each scenario, 100,000 packet arrivals are simulated and the average performances for the energy consumption and activation delay are derived.

Table 1: Costs of Energy Consumption and Activation Delay

State	Energy ($\mathcal{E}$)	Delay ($\mathcal{D}$)
DCH	100	0.15 s
FACH	40	0.48 s
PCH	2	1 s
IDLE	1	2 s

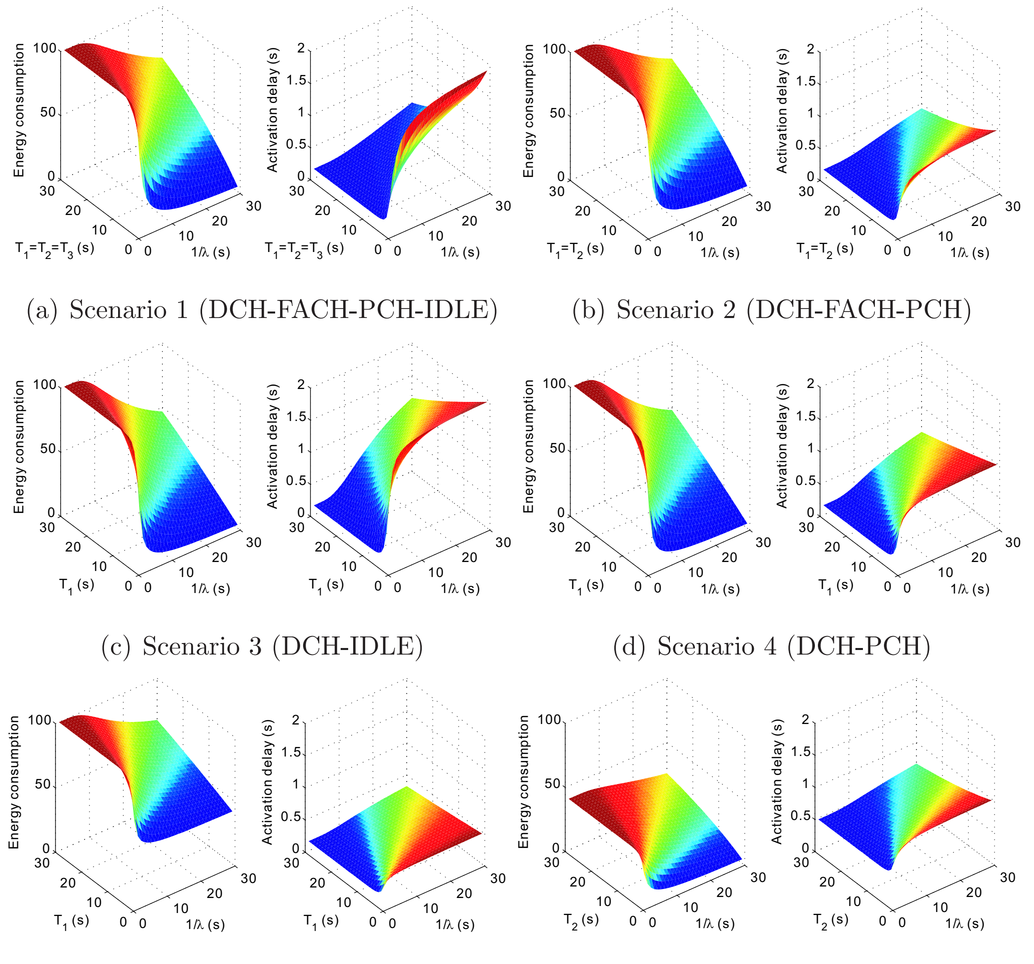
Fig. 5 shows the average energy consumption and activation delay as

Table 2: Considered State Management Scenarios

Scenario	T_1	T_2	T_3	Involved States
1	variable	variable	variable	DCH-FACH-PCH-IDLE
2	variable	variable	∞	DCH-FACH-PCH
3	variable	0	0	DCH-IDLE
4	variable	0	∞	DCH-PCH
5	variable	∞	n/a	DCH-FACH
6	0	variable	∞	FACH-PCH

functions of the inactivity timers and the average request packet inter-arrival time ($1/\lambda$) in the six scenarios. For the simplicity of parameter setting, we set the inactivity timers to be equal and vary them in case of scenarios 1 and 2. In all scenarios, the energy consumption increases and the activation delay decreases, as the inactivity timers increase and the inter-arrival time decreases. That is, the energy and delay have a tradeoff relationship with each other. As the inactivity timers increase and the inter-arrival time decreases, the UE is more likely to stay in the upper state (i.e., DCH or FACH state) rather than the lower state (i.e., PCH or IDLE state). This eventually increases the energy consumption but decreases the activation delay.

All scenarios show similar energy-delay tradeoff tendencies according to the timer values and the arrival rate, but their tradeoff performance ranges are different from each other. Comparing scenario 1 with scenario 2, the range of energy consumption is similar, but the range of activation delay is significantly decreased in the case of scenario 2. This is because scenario 2



(a) Scenario 1 (DCH-FACH-PCH-IDLE)

(b) Scenario 2 (DCH-FACH-PCH)

(c) Scenario 3 (DCH-IDLE)

(d) Scenario 4 (DCH-PCH)

(e) Scenario 5 (DCH-FACH)

(f) Scenario 6 (FACH-PCH)

Figure 5: Energy consumption and activation delay according to inactivity timers and packet inter-arrival time in the considered six scenarios.

excludes the IDLE state that has the highest delay cost, while accepting a small increase in energy consumption. In this context, scenario 5 using the DCH and FACH states shows the best performance in terms of the activation delay, and scenario 6 using the FACH and PCH states shows the best performance in terms of energy consumption.

Fig. 6 plots the energy-delay tradeoff curves of the six scenarios in the energy-delay plane when $1/\lambda$=10 s. Each curve is plotted by using (22)-(27)

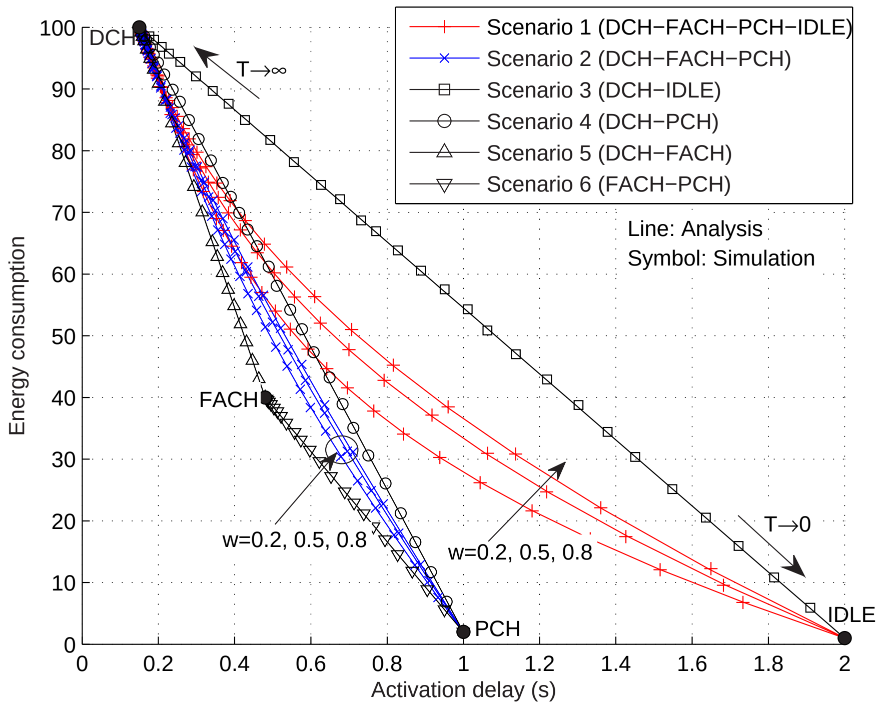


Figure 6: Energy-delay tradeoff curves.

and each black point denotes the energy-delay costs of each state. In all the scenarios, the energy consumption and the activation delay are inversely proportional to each other. More specifically, the energy-delay tradeoff curves in the scenario with three or four states (i.e., scenarios 1 and 2) are convexly decreasing. On the other hand, when only two states are involved, the tradeoff curve becomes a linear line so that the energy consumption and the activation delay are linearly inversely proportional to each other. The proof of this behavior and the numerical expression of this energy-delay tradeoff curve are provided in the Appendix.

Every point on each curve can be achieved by adjusting the inactivity timers. As the related timer value increases, the activation delay decreases but the energy consumption increases. Notice that an inner curve shows better tradeoff performances, i.e., both energy consumption and activation delay in the inner curve are smaller than those of the outer curve. The outer bound of the energy-delay tradeoff performance is given by scenario 3 (DCH-IDLE), and the inner bound of that consists of scenario 5 (DCH-FACH) and scenario 6 (FACH-PCH). The energy-delay tradeoff performances of the other scenarios are placed between these two bounds. It is also shown that the tradeoff curves of scenarios 1 and 2, which use more than two states, show convex form, and they are improved as the parameter w decreases (i.e., the buffer threshold increases). This implies that the use of FACH instead of DCH according to the downlink data traffic volume has a good effect on the tradeoff relationship.

The derived energy-delay tradeoff curve provides a guideline for network operators to suitably apply the user state transition mechanism according to the QoE requirements of various applications. For example, to reduce the activation delay, it would be better to use scenario 5. Similarly, the use of scenario 6 would be a better choice to reduce the energy consumption. More specifically, as shown in Fig. 6, when the activation delay requirement is less than 0.48 s, scenario 5 should be employed. Otherwise, it is better to use scenario 6 for saving energy. In practice, semi-real time services with tight delay constraints, such as stock update, are suitable for scenario 5. Non-real time services, such as in chatty applications, is more suitable for scenario 6. On the other hand, if we consider the channel utilization from a viewpoint of system capacity, it would be better to use scenario 2 (DCH-FACH-PCH), because it can flexibly allocate DCH or FACH according to the packet size in an efficient manner, even though the energy and delay performances are somewhat sacrificed. Additionally, if user mobility is considered, it would be desirable for high speed users to use scenario 1 including the IDLE state, because the IDLE state considerably reduces the signaling overhead caused by location update procedures due to its wide location update area.

5. Conclusion

In this paper, we analyzed the user state transition mechanism in terms of the energy consumption of UE and the uplink activation delay, and characterized them as a simple energy-delay tradeoff curve. This energy-delay tradeoff curve first informs the achievable performance bound of the user state transition mechanism, i.e., the tradeoff curve of the DCH-IDLE scenario shows the outer bound of the energy-delay tradeoff performance, and the tradeoff curves of both the DCH-FACH and FACH-PCH scenarios show the inner bound of that. In addition, the derived energy-delay tradeoff curve provides a guideline for network operators to suitably apply the user state transition mechanism according to the QoE requirements of various applications, i.e., the DCH-FACH scenario is suitable for reducing the delay and the FACH-PCH scenario is suitable for saving energy. For future study, we are planning to investigate another tradeoff issue considering the channel utilization and the signaling overhead from the perspective of network performance.

Appendix

Proposition 1. *When only two states are involved in the user state transi-*

tion mechanism, the energy-delay tradeoff curve becomes a linear line, i.e., the energy consumption and the activation delay are linearly inversely proportional to each other.

PROOF. Let us consider scenario 3 that uses only two states, DCH and IDLE. This scenario is made by setting the parameters as $T_2 = T_3 = 0$ and $w = 1$. By applying these values to (22)-(25), we obtain the steady state probabilities as follows:

$$\pi_{dch} = 1 - e^{-\lambda T_1}, \quad \pi_{fach} = \pi_{pch} = 0, \quad \pi_{idle} = e^{-\lambda T_1}. \tag{28}$$

As a result, from (26) and (27), the average energy consumption and the average activation delay are respectively represented as

$$\begin{aligned} \overline{\mathcal{E}} &= \pi_{dch}\mathcal{E}_{dch} + \pi_{idle}\mathcal{E}_{idle} = (1 - e^{-\lambda T_1})\mathcal{E}_{dch} + e^{-\lambda T_1}\mathcal{E}_{idle} && (29) \\ \overline{\mathcal{D}} &= \pi_{dch}\mathcal{D}_{dch} + \pi_{idle}\mathcal{D}_{idle} = (1 - e^{-\lambda T_1})\mathcal{D}_{dch} + e^{-\lambda T_1}\mathcal{D}_{idle}. && (30) \end{aligned}$$

Note that $(\overline{\mathcal{E}}, \overline{\mathcal{D}}) = (\mathcal{E}_{dch}, \mathcal{D}_{dch})$ as T_1 approaches infinity, and $(\overline{\mathcal{E}}, \overline{\mathcal{D}}) = (\mathcal{E}_{idle}, \mathcal{D}_{idle})$ as T_1 approaches zero.

The slope of the energy-delay tradeoff curve is defined according to the change of both the packet arrival rate λ and the inactivity timer T_1. By choosing arbitrary values, $(\lambda_1,\ T_1^1)$ and $(\lambda_2,\ T_1^2)$, the slope of the tradeoff curve is formulated as

$$\begin{aligned}
\frac{\Delta\overline{\mathcal{E}}}{\Delta\overline{\mathcal{D}}} &:= \frac{\overline{\mathcal{E}}_1 - \overline{\mathcal{E}}_2}{\overline{\mathcal{D}}_1 - \overline{\mathcal{D}}_2} \\
&= \frac{(1-e^{-\lambda_1 T_1^1})\mathcal{E}_{dch} + \ e^{-\lambda_1 T_1}\mathcal{E}_{idle} - \left\{(1-e^{-\lambda_2 T_1^2})\mathcal{E}_{dch} + e^{-\lambda_2 T_1^2}\mathcal{E}_{idle}\right\}}{(1-e^{-\lambda_1 T_1^1})\mathcal{D}_{dch} + \ e^{-\lambda_1 T_1}\mathcal{D}_{idle} - \left\{(1-e^{-\lambda_2 T_1^2})\mathcal{D}_{dch} + e^{-\lambda_2 T_1^2}\mathcal{D}_{idle}\right\}} \\
&= \frac{(e^{-\lambda_2 T_1^2} - e^{-\lambda_1 T_1^1})(\mathcal{E}_{dch} - \mathcal{E}_{idle})}{(e^{-\lambda_2 T_1^2} - e^{-\lambda_1 T_1^1})(\mathcal{D}_{dch} - \mathcal{D}_{idle})} \\
&= \frac{\mathcal{E}_{dch} - \mathcal{E}_{idle}}{\mathcal{D}_{dch} - \mathcal{D}_{idle}}. \qquad (31)
\end{aligned}$$

which is constant regardless of parameters λ and T_1. Therefore, it is verified that in scenario 3, the energy-delay tradeoff curve becomes a linear line connecting the energy-delay costs of DCH and IDLE states, $(\mathcal{E}_{dch}, \mathcal{D}_{dch})$ and $(\mathcal{E}_{idle}, \mathcal{D}_{idle})$, in the energy-delay plane. From this result, we can express the energy-delay tradeoff curve in scenario 3 as follow:

$$\overline{\mathcal{E}} = \frac{\mathcal{E}_{dch} - \mathcal{E}_{idle}}{\mathcal{D}_{dch} - \mathcal{D}_{idle}} \cdot \overline{\mathcal{D}} + \frac{\mathcal{E}_{idle}\mathcal{D}_{dch} - \mathcal{E}_{dch}\mathcal{D}_{idle}}{\mathcal{D}_{dch} - \mathcal{D}_{idle}} \qquad (32)$$

where $\mathcal{E}_{idle} \leq \overline{\mathcal{E}} \leq \mathcal{E}_{dch}$ and $\mathcal{D}_{dch} \leq \overline{\mathcal{D}} \leq \mathcal{D}_{idle}$.

Without loss of generality, a similar inducement can be applied to scenarios 4, 5 and 6, which ends the proof.

Acknowledgments

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (2011-0025424), and by the Human Resources Development program (No.20124030200060) of the Korea Institute of Energy Technology Evaluation and Planning (KETEP) grant funded by the Korea government Ministry of Trade, Industry and Energy.

References

3GPP TS 25.331. Radio Resource Control (RRC) protocol specification. 2012. v6.26.0.

Chiasserini C-F, Rao R-R. Improving Energy Saving in Wireless Systems by Using Dynamic Power Management. IEEE Trans. on Wireless Communications, 2003;2(5):1090-100.

Choi H, Lee J, Cho D. On the Use of a Power-Saving Mode for Mobile VoIP Devices and Its Performance Evaluation. IEEE Trans. on Consumer Electronics, 2009;55(3):1537-45.

Chung Y, Sung D, Aghvami A. Steady State Analysis of User Equipment State Transitions for Universal Mobile Telecommunication Systems. In: IEEE PIMRC, 2002. p. 2034-8.

GSM Association. Network Efficiency Task Force Fast Dormancy Best Practices. 2010. V1.0.

Holma H, Toskala A. WCDMA for UMTS: Radio Access for Third Generation Mobile Communications, 3rd Ed. John Wiley & Sons Ltd., 2004.

IEEE 802.16m-08/004r5. IEEE 802.16m Evaluation Methodology Document (EMD), 2009.

Jang S, Kim Y. Power Saving and Delay Reduction for Supporting WLAN-based Fixed-Mobile Convergence Service in SmartPhone. IEEE Transactions on Consumer Electronics, 2010;56(4):2747-55.

Lee J, Cho D. Performance Evaluation of Energy Saving Mechanism Based on Probabilistic Sleep Interval Decision in IEEE 802.16e. IEEE Trans. on Vehicular Technology, 2007;56(4):1773-80.

Nga D. T. T., Kim M, Kang M. Delay-guaranteed Energy Saving Algorithm for the Delay-sensitive Applications in IEEE 802.16e Systems. IEEE Trans. on Consumer Electronics, 2007;53(4):1339-47.

Perala P-H-J, Barbuzzi A, Boggia G, Pentikousis K. Theory and Practice of RRC State Transitions in UMTS Networks, In: IEEE GlOBECOM, 2009. p. 1-6.

Pries R, Magyari Z, Tran-Gia P. An HTTP web traffic model based on the top one million visited web pages. 8th EURO-NGI Conf. on Next Generation Internet (NGI), 2012. p. 133-9.

Ra M-R, Paek J, Sharma A-B. Energy-Delay Tradeoffs in Smartphone Applications. Conference on Mobile Systems, Applications, and Services (MobiSys'10), 2010.

Sarkar M, Cruz R.L. Analysis of Power Management for Energy and Delay Trade-off in a WLAN. Conference on Information Sciences and Systems, 2004.

Ukhanova A, Belyaev E, Wang Le, Forchhammer S. Power Consumption Analysis of Constant Bit Rate Video Transmission over 3G Networks, Elsevier Computer Communications, 2012;35(14):1695-706.

Vuyst S, Turck K, Fiems D, Wittevrongel S, Bruneel H. Delay versus energy consumption of the IEEE 802.16e sleep-mode mechanism, IEEE Trans. on Wireless Communications, 2009;8(11):5383-87.

Yeh J-H, Lee C-C, Chen J-C. Performance Analysis of Energy Consumption in 3GPP Networks. In: IEEE Wireless Telecom Symposium, 2004. p. 67-72.