DOI: 10.3745/KIPSTB.2008.15-B.6.585

HMIPv6에서 MN의 지연을 최소화하는 멀티미디어 컨텐츠 서비스 방법

박 원 길† · 강 의 선††

요 약

모바일 단말기를 사용하여 모바일 웹 서비스를 제공하는 과정에서 고려해야 할 사항은 끊김없는 서비스와 품질 보장형 서비스 이다. HMIIPv6는 끊김으로 인해 발생되는 패킷 손실과 지연 현상을 개선하기 위하여 MAP을 둔다. 하지만 MN을 위하여 패킷을 수신하고 전달하기 때문에 부하가 집중된다. 이는 보다 빨리 처리해야 할 실시간 데이터를 빨리 처리하지 못하는 결과는 낳는다. 뿐만 아니라 품질 보장 서비스를 위해서는 적응적 모바일 서비스 방법이 필요하다. 하지만 이 방법은 다양한 단말기의 하드웨어적인 차이점으로 인한 컨텐츠의 서비스 응답시간 비용이 필요하다. 따라서 본 논문에서는 끊김없는 서비스를 위하여 MAP에 큐를 두어 실시간 데이타 처리 성능을 향상시켰다. 또한 응답시간 비용을 줄이기 위하여 컨텐츠의 구성요소를 이용하여 컨텐츠를 재사용할 수 있는 캐시를 갖는 모바일 웹서비스 방법을 제안한다. 수식 분석과 시뮬레이션 결과, 제안방법인MAP에 큐를 두는 방식과 모바일 노드의 이동성을 고려한 캐시 방법이 다른 다양한 시스템 조건에서 우수함을 알 수 있었다.

키워드 : HMIPv6, 핸드오버, 멀티레벨 큐, 모바일 웹 서비스, 적응화, 캐쉬

The Adaptive Multimedia Contents Service Method to Reduce Delay of MN in HMIPv6

Park, Wongil† · Kang, Euisun††

ABSTRACT

The issues that we should consider in the process of providing mobile web service using a mobile device are seamless service and QoS-guaranteed service. HMIPv6 has MAP because of improving packet loss and transmission delay due to disconnection. However, a load is concentrated on HMIPv6 because of receiving and delivering packet for MN. Owing to this, real time data fails to be processed quickly, and also adaptive mobile service is required for QoS guaranteed service. However, this method demands the response time cost of contents service owing to the hardware differences of various devices. Therefore, we improve the process performance of real time data by applying a queue in MAP for seamless service in this paper. For decreasing response time cost, we propose mobile web service method which has reusable cache of contents using the elements of contents. The result of a numerical formula and simulation shows that our proposed method is superior under various system conditions.

Keywords : HMIPv6, Handover, Multi-Level Queue, Mobile Web Service, Adaptation, Cache

1. Introduction

Communication service all over the world is developed through the paradigm such as broadband, multimedia, wired/wireless integration and service unification. Therefore, the future society will enjoy various seamless QoS guarantee Internet services regardless of device type anywhere and anytime. For providing seamless service, mobility should be considered and the adaptive service method according to the environment of device for QoS-guarantee service. Unlike wired environment, a device having mobility should always be possible to communicate through the Internet even if it is moving in wireless environment. For this, MIPv6(Mobile IPv6) is designed MIPv6 is the expansion of existing IP protocol for maintaining connection although a mobile node that is a mobility host, change a point-of-attachment[1]. MIPv6 is MN

※ 본 연구는 숭실대학교 교내연구비 지원으로 이루어졌음.
† 정 회 원 : 한국건설기술연구원 박사후 연구원
†† 정 회 원 : 숭실대학교 교양특성화대학 전임강사(교신저자)
논문접수 : 2008년 7월 15일
수 정 일 : 1차 2008년 9월 11일
심사완료 : 2008년 9월 16일

(Mobile Node), CN (Correspondent Node) which communicate with MN, and HA (Home Agent) which is the home network router of MN. In case the distance between two hosts moving in MIPv6, the process time for binding update is long, thus the traffic loss and transmission delay will occur. For solving these problems, HMIPv6 (Hierarchical Mobile IPv6) emerged. HMIPv6 introduces an agent called MAP (Mobility Anchor Point) for solving the binding problem of MIPv6. MAP acts as HA and receives the transmitted packets to the location serviced by it instead of MN and then transmits to the current location of MN after encapsulation. Subsequently, it processes the transmitted data for MN. Therefore, a problem related to load concentrates on MAP arises. The concentration problem can affect whole network, which requires a solution.

When MN moves between MAP, CN is possible to communicate with MN without passing HA after binding update. At this time, CN which provides web service is confronted by the problems of creating and servicing existing web contents. Therefore, for providing same contents to MN having various hardware specifications, the contents adaptation considering the environment of MN is required. The content adaptation methods are converting methods [2] as a language that MN can understand, real time service method [3-4] which services multimedia data on web to MN. However, the proposed methods contain real time or not real time conversion time for adaptation so we should consider time cost. Therefore, the quantity of mobile contents and the kind of device have a strong influence on the response time of service according to the conversion time and also the server capacity to store various versions of contents. Finally, we need a method to transmit Internet service continually and simultaneously to decrease the handover for providing same contents to various devices. Besides, we also need a method for decreasing response time occurred when CN provides adaptively to various devices, and the server capacity.

2. Related Works

HMIPv6 was proposed to solve local mobility management problem of MIPv6. HMIPv6 is used as a MAP. There are several issues relating MAPs. The first issue is on the load management of MAPs. It depends on the MAP configuration and the number of MN. The second issue is to solve relation establishment for movement characteristic of MN of special situation and MAP relation. For example, it is a problem of situation that a MAP concentrates on MN in a special time zone or a special place. Because MAP act as local HAs, it provides better performance. If MN moves in MAP domain (micro handover), the performance is within our expectation because MN register binding update message to MAP. However, if MN moves MAP and MAP (macro handover), MN is not only register binding update message to MAP but also HA or CN. It happens to add binding update message. If the movement speed of MN is fast, macro handover often happens. It can reduce performance of system [5-7]. Most of studies related to MIPv6 and HMIPv6 aim to reduce handover delay which causes packet loss when MN moves faster, or to reduce signaling method on transmission to MN and HA, or the time MN spends during register operation process to exchange message when it registers binding update [8-12].

To reduce response time, there are several caching methods considering the mobility of devices in adaptive mobile web service. Mobile devices have widely different characteristics from PC such as mobility, limited network environment and battery, frequent access seam and small screen size. Specially, it is more important to manage web cache in mobile. There are two methods for mobility of portable device. One is semantic cache method that uses movement of device [13-15]. The other is cache method that uses available range of space object [16]. Semantic cache method [13-15] is to utilize local position of user and to set individual service area based on direction and speed of user or service content. And then it replaces objects in cache and sends suitable information to mobile according to the area. That objects are saved within available range [16] in cache focusing on how to decide the area of object. Firstly, this method represents available range by a circle with center point and a radius of the object. And then objects in the area are stored in cache. To consider movement of device, the methods focus on how to accurately transmit some objects according to mobility and speed. Original web caches are divided according to physical position of cache, server side and client side. However, environment of MN is limited than PC. This means that it is almost impossible to have cache inside MN such as cellular phone. Therefore, we apply web cache method between server and MN using temporal characteristic. In real-time, there are three methods. One is based on reference time [20], another is based on reference frequency [21] and the other is base on object size [22]. The algorithms set priority order of web document in order to store in limited cache. Accordingly, this paper introduces some methods. Some methods which are able to transmit multimedia object seamlessly and also to reduce handover from mobility of MN in case single content is serviced to various device are needed.

3. MAP Management Methods using Multilevel Queue

HMIPv6 is using MAP of new agent; its role is on HA instead of MN. MN incoming MAP Domain receives RA message to have information local MAP more than one. MN can be bound to present location of oneself to have RCoA include subnet of MAP. The MAP role of HA receives all packets for MN, and it transmits present address of MN. Therefore if MN moves in MAP area, it only needs to register its address to MAP. However, if MN moves another MAP, it needs to register with not only MAP but HA or CN. HMIPv6 is proposed to reduce signaling overhead and handover of MIPv6. MAN may have concentration because MAP of HMIPv6 receives all packets instead of MN, it forwards the present location of MN. These concentrations have influence on the whole network, so we proposed that the MAP have multilevel queue in order to process binding update message. We add one bit existing binding update message and we call P(Priority)bit. MAP is different according to P bit of binding update message. It depends on the processing data, whether it is real time data or non real time data, or interactive and non interactive. We show priority process of MAP in (Fig. 1).

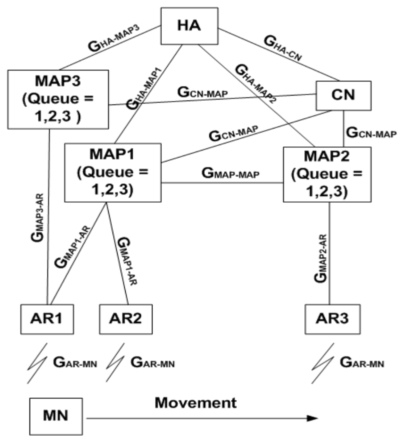
(Fig. 1) Three level queue of MAP

3.1 Cost Functions

3.1.1 The Cost of Location Update

(Fig. 2) shows proposed system model for analysis cost of proposed system.

In HMIPv6, a MN performs two types of binding update procedure: the macro binding update and the micro binding update. MN registers its RCoA with the CNs and the HA in macro binding update. On the other hand, if a MN changes its current address within a local MAP



(Fig. 2) Network Model

domain. We divide when MN moves between networks, the total cost into location update cost and packet delivery cost in HMIPv6. Therefore, the total cost is as follows:

$$C_T = C_L + C_P \tag{1}$$

$C_{L-global}$, $C_{L-new-global}$, $C_{L-local}$ and $C_{L-new-local}$ denote the signaling costs in the global binding update, the global binding update of proposed scheme, the local binding update and the local binding update of proposed scheme, respectively.

$$\begin{aligned} C_{L-global} &= 2\cdot(k\cdot G_{AR-MN} + \tau\cdot(G_{HA-MAP\{1,2\}} + G_{MAP\{1,2\}-AR})) \\ &+ 2\cdot N_{CN}\cdot(k\cdot G_{AR-MN} + \tau\cdot(G_{HA-MAP\{1,2\}}) + G_{CN-MAP})) \\ &+ PC_{HA} + N_{CN}\cdot PC_{CN} + 2\cdot PC_{MAP} \end{aligned} \tag{2}$$

$$\begin{aligned} C_{L-new-global} &= 2\cdot(k\cdot G_{AR-MN} + \tau\cdot(G_{HA-MAP3} + G_{MAP3-AR})) \\ &+ 2\cdot N_{CN}\cdot(k\cdot G_{AR-MN} + \tau\cdot(G_{HA-MAP3} + G_{CN-MAP})) \\ &+ PC_{HA} + N_{CN}\cdot PC_{CN} + PC_{MAP} \end{aligned} \tag{3}$$

$$C_{L-local} = 2\cdot(k\cdot G_{AR-MN} + \tau\cdot G_{MAP\{1,2\}-AR}) + 2\cdot PC_{MAP} \tag{4}$$

$$C_{L-new-local} = 2\cdot(k\cdot G_{AR-MN} + \tau\cdot G_{MAP3-AR}) + 2\cdot PC_{MAP} \tag{5}$$

Here τ and k are the unit transmission costs in a wired and wireless link, respectively. Let G_{HA-MAP}, G_{MAP-AR}, G_{AR-MN} and G_{CN-MAP} are the hop distance between nodes. It is proportional hop number. PC_{HA}, PC_{CN} and PC_{MAP} are the processing costs for binding update procedures at the HA, the CN and the MAP, respectively. N_{CN} denotes the number of CNs which is communicating with the MN. In terms of the random walk mobility model, the probability that a MN performs a macro handover is as follows:

$$P_{D,D}\cdot\alpha_{d,d+1}$$

Specifically, if a MN is located in range ring D, the boundary ring of a MAP domain is composed of D range rings, and performs a movement from range ring Dto range ring D+1. The MN then performs the macro handover procedure.

In other cases, except this movement, the MN only performs a micro handover procedure. Hence, the location update cost of existing and proposed scheme per unit time can be expressed as follows:

$$C_L = \frac{P_{D,D} \cdot \alpha_{D,D+1} \cdot C_{L-global} + (1 - P_{D,D} \cdot \alpha_{D,D+1}) \cdot C_{L-local}}{T} \quad (6)$$

$$C_{new-L} = \frac{P_{D,D} \cdot \alpha_{D,D+1} \cdot C_{L-new-global} + (1 - P_{D,D} \cdot \alpha_{D,D+1}) \cdot C_{L-new-local}}{T} \quad (7)$$

3.1.2 The cost of packet delivery

The packet delivery cost, C_P, in HMIPv6 can then be calculated as follows:

$$C_P = C_{P-MAP} + C_{P-ha} + C_{P-cn-mn} \quad (8)$$

In (Eq. 8), C_{P-MAP} and C_{P-ha} denote the processing costs for packet delivery at the MAP and the HA, respectively. $C_{P-cn-mn}$ denotes the packet transmission cost from the CN to the MN. The cost from CN to MN of C_P can be expressed as follows:

$$\begin{aligned} C_{P-cn-mn} &= \tau \cdot \lambda \cdot ((S-1) \cdot (G_{CN-MAP} + G_{MAP\{1,2\}-AR}) \\ &+ (G_{HA-CN} + G_{HA-MAP\{1,2\}} + G_{MAP\{1,2\}-AR})) + k \cdot \lambda \cdot S \end{aligned} \quad (9)$$

$$\begin{aligned} C_{P-new-cn-mn} &= \tau \cdot \lambda \cdot ((S-1) \cdot (G_{CN-MAP} + G_{MAP3-AR}) \\ &+ (G_{HA-CN} + G_{HA-MAP3} + G_{MAP3-AR})) + k \cdot \lambda \cdot S \end{aligned} \quad (10)$$

The packet transmission cost according to binding update message of multilevel queue to decentralization load at MAP can be expressed as follow:

$$\begin{aligned} C_{P-high} &= C_{P-MAP-high} + C_{P-ha} + C_{P-new-cn-mn} \\ C_{P-mid} &= C_{P-MAP-mid} + C_{P-ha} + C_{P-new-cn-mn} \\ C_{P-low} &= C_{P-MAP-low} + C_{P-ha} + C_{P-new-cn-mn} \end{aligned} \quad (11)$$

In (Eq. 11), the packet delivery cost of C_{P-high}, C_{P-mid} and C_{P-low} can be calculated as follows:

$$C_{P-ha} = \lambda \cdot \theta_{HA} \quad (12)$$

Where Θ_{HA} refers to a unit packet processing cost at the HA. Applying FIFO to three kinds of traffics according to P bit values such as 0, 1, 2 we can get M/G/1.

$$\begin{aligned} &\lambda = \lambda_1 + \lambda_2 + \lambda_3, \\ &G(x) = \frac{\lambda_1}{\lambda} G_1(x) + \frac{\lambda_2}{\lambda} G_2(x) + \frac{\lambda_3}{\lambda} G_3(x) \end{aligned} \quad (13)$$

3 types of traffic such as high class, medium class, low class according to Poisson process of parameter λ_1, λ_2, λ_3 have service distribution of G_1, G2 and G3. By this rule, we can compute as follows [13], which follow since the combination of two independent Poisson process is itself a Poisson process whose rate is the sum of the rates of the component processes. By this rule, we can compute W_Q^i as follows [13].

$$\begin{aligned} W_Q^1 &= \frac{\lambda_1 E[S_1^2] + \lambda_2 E[S_2^2] + \lambda_3 E[S_3^2]}{2(1 - \lambda_1 E[S_1])} \\ W_Q^2 &= \frac{\lambda_1 E[S_1^2] + \lambda_2 E[S_2^2] + \lambda_3 E[S_3^2]}{2(1 - \lambda_1 E[S_1])(1 - \lambda_1 E[S_1] - \lambda_2 E[S_2])} \\ W_Q^3 &= \frac{\lambda_1 E[S_1^2] + \lambda_2 E[S_2^2] + \lambda_3 E[S_3^2]}{2(1 - \lambda_1 E[S_1] - \lambda_2 E[S_2])(1 - \lambda_1 E[S_1] - \lambda_2 E[S_2] - \lambda_3 E[S_3])} \end{aligned} \quad (14)$$

Here, λ_1, λ_2, λ_3 is arrival rate per time of high class, medium class, low class customer. $E[S_1]$ is that high class customer wait average waiting time for service in queue. $E[S_2]$ is that medium class customer wait average waiting time for service in queue. $E[S_3]$ is that low class customer wait average waiting time for service in queue.

This paper assumes that the average number of users located in the coverage of an AR is k. Therefore, the total number of users can be obtained using (Eq. 15).

$$N_{MN} N_{AR} \times k \quad (15)$$

Consider the information above, packet treatment cost (C_{P-MAP}) in MAP, packet treatment cost ($C_{P-MAP-high}$) in high class of MAP, packet treatment cost ($C_{P-MAP-mid}$) in mid class of MAP, packet treatment cost ($C_{P-MAP-low}$) is can be calculated as follow.

$$\begin{aligned} C_{P-MAP} &= \lambda \cdot S \cdot (C_{lookup} + C_{routing} + C_{wait}) \\ &= \lambda \cdot S \cdot (\alpha \cdot N_{MN} + \beta \log(N_{AR}) + W_Q) \end{aligned} \quad (16)$$

$$\begin{aligned} C_{P-MAP-high} &= \lambda \cdot S \cdot (C_{lookup} + C_{routing} + C_{wait}) \\ &= \lambda \cdot S \cdot (\alpha \cdot N_{MN} + \beta \log(N_{AR}) + \delta \cdot W_Q^1) \end{aligned} \quad (17)$$

$$C_{P-MAP-mid} = \lambda \cdot S \cdot (C_{lookup} + C_{routing} + C_{wait})$$
$$= \lambda \cdot S \cdot (\alpha \cdot N_{MN} + \beta \log(N_{AR}) + \delta \cdot W_Q^2) \quad (18)$$

$$C_{P-MAP-low} = \lambda \cdot S \cdot (C_{lookup} + C_{routing} + C_{wait})$$
$$= \lambda \cdot S \cdot (\alpha \cdot N_{MN} + \beta \log(N_{AR}) + \delta \cdot W_Q^3) \quad (19)$$

In (Eq. 16, 17, 18 and 19) λ denotes the session arrival rate and S denotes the average session size in the unit of packet. α, β and δ are the weighting factors. We show some evaluation results to demonstrate the performance in chapter 5.

4. Pre-Service Method based on Content

4.1 Pre-Service Method using re-using content in Cache

In general, web caching aims to enhance response time to the repetitively serviced page by storing user contents. However, despite using it, creating contents to suit characteristic of diverse terminal environment in mobile service may bring to make multi-versions for single content. It demands frequent replacing of the cache due to multi-versions. CN would acts as server or gateway for mobile web service that repeatedly replacing cache is to require converting content to servicer. Therefore, even with cache application, a method is required to reduce time taken to create or transcode mobile contents and to minimize response time. In order to improve such weak-nesses, this paper proposes the method of Pre-Service.

Let us assume that there is a mobile phone with resolution of 120×96. The size of maximum image to be displayed in this device would be 120×96 in case vertical and horizontal scroll bars does not have. Therefore, image contents should be converted and transmitted to suit the size of 120×96 by using markup language. However, service is possible without any transcoding in terminals that have the same markup language and resolution (ex:176×144) higher than 120×96. The image will be viewed small in the eyes of users.

Pre-Service is to choose playable content in terminal among a number of contents in cache and to service them first even if contents of the optimal quality cannot be serviced in case the contents to suit the requesting terminal do not exist in cache. Here, the playable mobile page refers to the range in which each object of the page is visible in browser of the requesting MN even though the selected mobile page may be of low quality. Selecting is based on threshold values, the range of possible service. These threshold values are determined by user or server administrator in range of every playable object in page.

Pseudo Code for Algorithm of Pre-Service

```
Pre-Service Algorithm
-----------------------------------------------------------
md : requested content by MN
mci : ith content in Cache
T : Threshold
TR : Transcoding mci
Diff(mci, md) : difference degree between mci and md
flag : success flag for searching playable content in Cache

Procedure Pre-Service
Get HTTP_USER_AGENT from  MN
Get PageID from md
DeviceQosList = Analyzing(HTTP_USER_AGENT)

count = CountContent(Cache)
 for i=1 to count
Get ID from mci
        if PageID == ID then
             calculate Diff(mci, md)
             if Diff(mci, md) ≥ T then
                flag = true
                break;
             else
                flag = false
            endif
endfor

if flag ≠ true then
     mci = TR(mci)
     Replace Cache(mci, QosList)
endif
     Response mci to Device
```

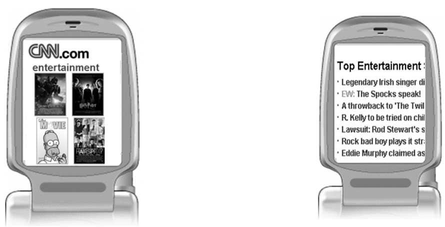
Firstly, CN waits for service requirement for specific content from a MN. When a specific terminal has been connected, it gets HTTP_USER_AGENT of the terminal to extract terminal information and analyze the contents requested. Secondly, it searches for the mobile contents to suit the service-requesting terminal in cache. At this time, it compares requested content by MN to content in cache using selecting method described in Section 4.3. In cache, if there is content that either exactly corresponds or displays on browser in accessed MN, CN sends the content and waits for the next request from MN.

4.2 Content aware selecting method

CN may need to select content that is similar to access MN in order to provide faster service by reusing content in cache. Here, similar content is both to be able to display content on MN and to have the same object elements in page at once. In other words, it selects content that the difference of elements information for object in the page is least between request content and content in cache. In this section, we suppose a method of content aware selecting by choosing playable content of multitude version. In general, a single mobile web page consists of a multiple number of objects expressible by markup language.

As shown above, (Fig. 3) is the mobile contents consist of images and texts. As (Fig. 3), “Content A” needs Qos (Quality of Service) Information for Image and “Content B” needs them for text. More detail, to present “Content A” of (Fig. 3) in MN, QoS information of each object is required such as markup language, image type, color depth and resolution. Therefore, content aware selection method is where objects to configure a single mobile web page use the configuration elements. It is where objects to configure a mobile web page are analyzed, only the information required in expressing the objects is extracted and compared. Mobile contents create Markup language on the basis of QoS information of the connected terminal. Therefore, it is possible to extract essential elements required for displaying from contents.

For example, let assume that horizontal size of accessed MN is 240 and the size of C content in cache is 128. The rate of their width resolution is 0.53(=128/240). Namely,



<Content A> < Content B>

(Fig. 3) Example of object element in mobile content

〈Table 1〉 Example of QoS Information for MN and content in Cache

	QoS	Express (Numerical)	QoS of MN	QoS of Content in Cache		
				A	B	C
1	MarkUp	1, 0	mHTML	WML	mHTML	mHTML
2	width	Num	240	240	240	128
3	height	Num	320	320	320	160
4	Color Depth	Num	4096	256	256	4096
5	Image Format	1, 0	JPG	JPG	JPG	GIF
6	Sound Poly	Num	64	4	40	16
7	Melody Format	1, 0	MA	MF	MA	MSND
8	Video Format	1, 0	MPEG4	MPEG4	MPEG4	H.263

53% means the difference degree of horizontal resolution.

For example, request device has mHTML as markup langue and some content in cache has same langue. When CN compares their markup, it is expressed as 1 when languages of contents are the same and as 0 when not. Markup language differs according to protocol or browser used in MN. Therefore, service itself cannot be possible in case serviceable markup language is not supported in terminal.

〈Table 2〉 Symbols to calculate the difference between content

symbol	meaning
mc_i	the ‘i’th mobile contents saved in cache
md	the MN to request service
QoS_j	the QoS information of ‘j’
O_i	each object obtained by parsing mobile page.
QoS_f	Multimedia data format of object(O_i)
QoS_q	extra information of object(O_i) to more enhance service quality
QoS_{Markup}	Markup language in mc_i or md

To calculate contents similarity, the following preconditions are required.

$$QoS_j(mc_i) \le QoS_j(md) \tag{20}$$

The reason for precondition as shown in (Eq. 20) is that it is impossible for service contents to hold QoS information greater than QoS information of the service-requesting terminal. Even if it were possible, accurate information delivery would be unattainable. The similarity of contents is calculated by using the following.

$$mc_i \in O$$

$$O = \{O_1, O_2, O_3, \ldots, O_n\}$$

$$O_i = \{< QoS_f, QoS_q > | QoS_f \in Format, Qos_q \in ExInfo\} \tag{21}$$

In (Eq. 21), O means the group of objects in mobile page. mc_i has ‘n’ number of multimedia data objects. Each object in page is obtained by farcing mobile page. O_i signifies each object obtained by parsing mobile page. Each object of mobile page consists of multimedia format and additional information. The (Eq.22) represents the computed result value that it compares information of each element on content with requested content in order to find out content in Cache.

$$QoS_{Markup}(mc_i, md) = \begin{cases} 1 & QoS_{Markup}(mc_i) = QoS_{Markup}(md) \\ 0 & QoS_{Markup}(mc_i) \neq QoS_{Markup}(md) \end{cases}$$

$$QoS_f(O_i) = \begin{cases} 1 & Format_{mc_i}(O_i) = Format_{md}(O_i) \\ 0 & Format_{mc_i}(O_i) \neq Format_{md}(O_i) \end{cases} \quad (22)$$

$$QoS_q(O_i) = \frac{ExInfo_{mc_i}(O_i)}{ExInfo_{md}(O_i)} \quad ,0 < QoS_q(O_i) < 1$$

Finally, the difference between requested content and special content is calculated by (Eq. 23) and a content in cache is selected as the most similar it. The content is that either the result has the biggest value or the difference is the smallest value.

$$Diff(mc_i, md) = QoS_{Markup}(mc_i, md) \times (\omega_i \times \sum QoS_f(O_i) + \omega_j \times \sum QoS_q(O_i))) \quad (23)$$

Ultimately, $Diff(mc_i, md)$ with application of ω_i and ω_j, signifies the calculated similarity of contents. In process of $Diff(mc_i, md)$, ω_i should be bigger than ω_j because information of multimedia format is more important than it of addition. For Example, while image may be shown a little on browser in case that image size of component in page is bigger than browser's size on MN, the wrong and erroneous data may be presented to MN in case the multimedia format is different.

5. Simulation Results

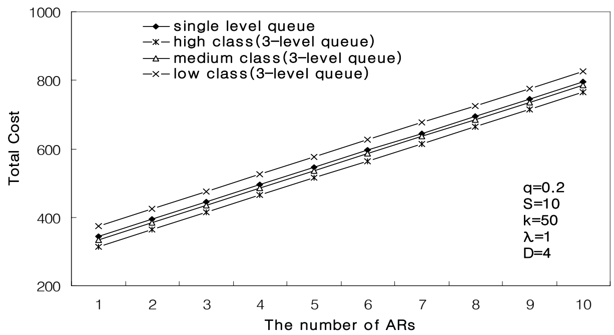
In this section, we provide some numerical evaluation to demonstrate the performance of proposed scheme as compared with normal HMIPv6. In general, to enhance response time using cache has relationship to hit ratio. To find document and service in cache without trans-coding contents in the course of service means that the time required in transcoding contents decreases. Therefore, performance evaluation was administered on the basis of cache hit ratio and the result was compared for analysis. The experiment and test were administered in Windows XP Professional with the specifications of 2.66GHz and 1GB memory The parameter values for the analysis were referenced from [17-19], as shown in <Table 3>.

5.1 MAP management scheme using multilevel queue

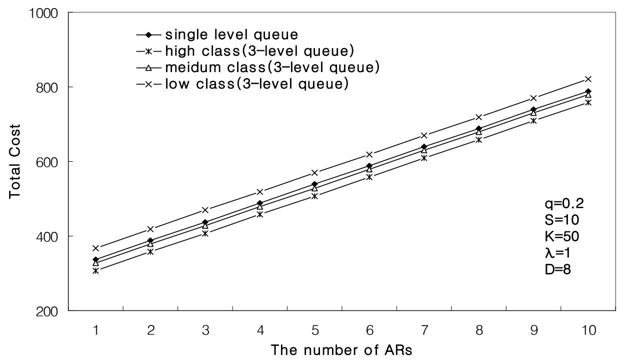
(Fig. 4) and (Fig. 5) show the value of total cost as the number of AR. As shown in (Fig. 1) and (Fig. 2), the total cost increases linearly as the number of AR increases. With the multilevel queues, we reduce the total

〈Table 3〉 Numerical simulation parameter for performance analysis

Parameter for performance analysis of multilevel queue			
parameter	value	parameter	value
α	0.1	G_{HA-CN}	6
β	0.2	G_{CN-MAP}	4
ɣ	0.05	$G_{MAP-MAP}$	1
σ	10	G_{MAP-AR}	2
θ_{HA}	20	G_{AR-MN}	1
τ	1	PC_{CN}	6
k	2	PC_{HA}	24
λ_1	0.1	PC_{MAP}	12
λ_2	0.2	G_{HA-MAP}	6
N_{cn}	2		
Parameters used in CN			
The number of page	20	the number of access pattern	5
Objects in Page	Image, Text	replace algorithm	LFU
the number of test device	10	threshold	100%, 90%, 80%, 70%, 60%
cache capacity	100KB	the number of access user	100, 200, 300, 400, 500



(Fig. 4) The value of total cost as function of the number of AR(D=4)

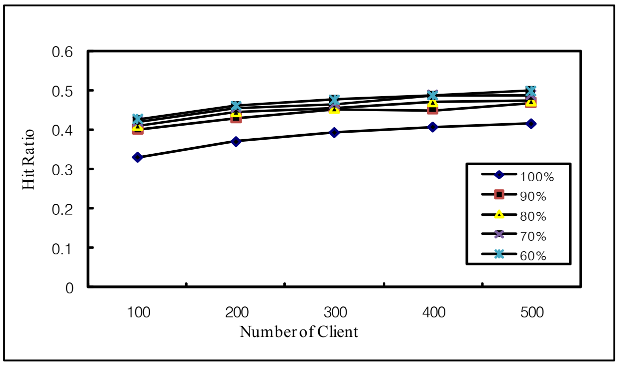


(Fig. 5) The value of total cost as function of the number of AR(D=8) cost by 10% and 12% when the size of D is 4 and 8 in high class respectively. In comparison of proposed scheme with existing method, our priority method reduces the total cost by 12% approximately. The cost when we use multilevel queue with priority shows superior performance in real time process rather than does not use multilevel queue.

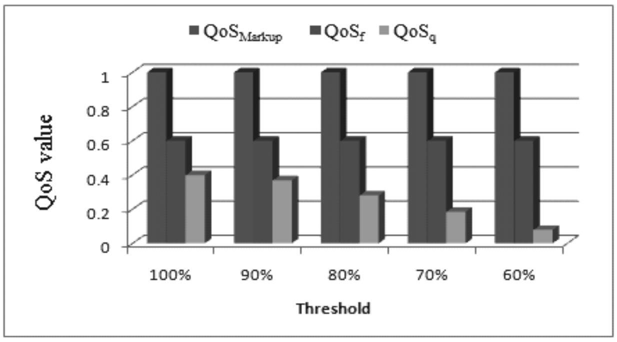
5.2 The performance about content selecting method in CN

(Fig. 6) shows hit ratio in cache when CN performs content selection method. And (Fig. 7) represents the result of each QoS according to threshold in case of applying to pre-service.

In the figures above, 100%~60% mean threshold is assigned by user. As shown by (Fig. 6), cache hit ratio improved by approximately 10% when serving contents of 90~60% similarity with slightly low level of quality service rather than when servicing contents of 100% similarity. Because there are more content about 90%~60% than exactly corresponding to request content. (Fig. 7) shows the result performed according to content information in case server manager set threshold. The result values are to calculate averages about information selected contents within threshold range by pre-service. As shown in figure, QoS_{Markup} and QoS_f only satisfy in case of equal to requested content due to determine whether the content can be serviced or not. Therefore, at the point of choosing similar content in cache with request content, QoS_q is important element and decide the rest of accuracy. In this paper, we assign that ω_i is 0.6 and ω_j is 0.4 for weight values. The reason is that every object can be serviced better in playable range. If the server manager considers fast response time, he/she should set higher ω_j. However, quality should be reduced.



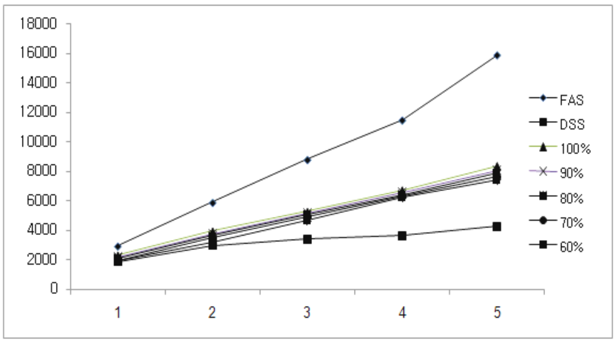
(Fig. 6) Hit ratio according to number of Client



(Fig. 7) QoS value according to threshold

The methods of servicing mobile web pages created to diverse mobile terminals [23] are divided into device-specific service(DSS) to draw up each content in advance to suit characteristics of terminals and full-automatic service(FAS) to automatically recreate contents to suit characteristics of terminals when service request is received from client.

In (Fig. 8), when server receives a request from MN, FAS is to service automatically re-created content to MN. And DSS is to send content to MN. Content is previously transformed according to hardware characteristic of MN. FAS need cache because of real-time transcoding. This paper use LFU(Least frequency Used) of reference frequency cache methods. LFU always keeps the most frequently used items in cache As shown in (Fig. 8), response time displayed overall improvement from that of full-automatic service when using cache[20]. It decreased by 10% when the count of persons logged on was 100 and by 25% when the count was 500. Although it drops by 25% when the number of persons log on was 500, improvement of performance can be anticipated in case more people log on to the server. Unlike pre-service, full-automatic service method is where terminal services content after transcoding them at the time of request on real time basis. Although the time of transcoding contents is short, it indicates the importance of cache as the number of persons simultaneously logs on increases. Furthermore, it indicates that the method of transcoding after pre-service produces superior performance.



(Fig. 8) Content based selecting Method

6. Conclusion

Since the rapid growth of Internet, the development of wireless network makes the demand of Internet service increase in mobile communication environment. As such, HMIPv6 is proposed to guarantee the mobility of device and various methods are introduced for applying existing Internet schemes to mobile. HMIPv6 introduces MAP, which is an agent that acts as HA(Home Agent) for mobile node, so much load is concentrated. The real time data cannot be quickly processed than non real time because of load concentration. Therefore, this paper proposes the method which can process real time data faster by putting queue on MAP. For the mobile service's side, it has problem that server-side should executes frequent conversion for mobile node having various hardware characteristics. In case the contents requested by a device do not exist in a cache, we propose the method that selects the replaying contents among several contents instead of servicing the most suitable contents within the cache and then services ahead. From the result of numerical expression and simulation, we may conclude that the method which put a queue on MAP and the cache method considering the mobility of MN are superior under different conditions of various systems.

Reference

[1] D. Johnson, C. Perkins and J. Arrko, "Mobility Support in IPv6," Internet Draft, IETF, draft-ietf-mobileip-ipv6-20.txt(work in progress), Jan., 2003.

[2] E. Kirda, C. Kerer and G. Matzka, "Using XML/XSL to build adaptable database interfaces for Web site content management," XML in Software Engineering Workshop, 23rd International Conference on Software Engineering, Toronto, Ontario, Canada, 2001.

[3] M. Hong, D.H.k Park, Y.H. Lim., "Transcoding Pattern Generation for Adaptation of Digital items Containing Multiple Media Streams inUbiquitous Environment," LNCS 3583, pp.1036-1045, 2005.

[4] E.S Kang, M Hong, Y.H Lim., "A Guided Search method for Real time Transcoding a MPEG2 P frame into H.263 P Frame in a Compressed Domain," LNCS 3581, pp.242-251, 2005.

[5] H. Soliman, C. Castelluccia, K. E. Malki and L, Bellier, "Hierarchical MIPv6 Mobility Management (HMIPv6)," Internet Draft, Nov., 2001.

[6] R. Hsieh, Zhe Guang Zhou and Aruna Seneviratne, "S-MIP: A Seamless Handoff Architecture for Mobile IP," Proceedings of INFOCOM 2003, San Francisco, pp.1774-1784, April, 2003.

[7] R. Hsieh, A. Seneviratne, H. Soliman and K. El-Malki, "Performance Analysis on Hierarchical Mobile IPv6 with Fast-handoff over end-to-end TCP," Proceedings of GLOBECOM 2002, Taipei, Taiwan, Vol.3, pp.2488-2492, Nov., 2002.

[8] I. Vivaldi, M.H. Habaebi, B.M. Ali and V. Prakesh: "Fast Handover Algorithm for Hierarchical Mobile IPv6 Macro-mobility Management," Association of Professional Communication Consultants 2003, Penang Malaysia, Vol.2, pp.630-634, September, 2003.

[9] B. Shin and K. Ha, "An Improvement of Handoff Latency by Virtual MAPs for HMIPv6," Proceeding of Conference on Information Technology 2006, Bhubaneswar India, p.78, Dec., 2006.

[10] J. Jang, D. Kown and Y. Suh, "An Efficient Handoff Mechanism with Reduced Latency in Hierarchical Mobile IPv6," Proceedings of International Conference on Information Networking, Busan Korea, pp.184-194, February, 2004.

[11] W. Zheng, P.X. Zeng, C. Jian and C.Y. Zeng, "A three-level Mobility Management Scheme for Hierarchical Mobile IPv6 Networks," Journal on Zhejiang University SCIENCE A, July, 2006.

[12] Y. W. Chen and M. J. Huang, "A Novel MAP Selection Scheme by Using Abstraction Node in Hierarchical MIPv6," Proceeding of IEEE International Conference on Communications, Istanbul Turkey, Vol.12, pp.5408-5413, Jun., 2006.

[13] Q. Ren and M.H. Dunhanm, "Using Semantic Caching to Manage Location Dependent Data Mobile Computing," In Proceedings of MobilCom, pp.210-221, 2000.

[14] R. Lee, K. Goshima, Y. Kambayashi and H. Takakura, "Caching Schema for Mobile Web information Retrieval," 2nd International Workshop on Web Dynamics, in conjunction with WWW, 2002.

[15] A. Pashtan, A. Heusser and P. Scheuermann, "Personal Service Areas for Mobile Web Applications," IEEE Internet Computing Vol.8, No.6, pp.34-39, Nov./Dec., 2004.

[16] B. Zheng, J. Xu and D. L.Lee, "Cache Invalidation and Replacement Strategies for Location-Dependent Data in Mobile Environment," IEEE Transactions on Computers, Vol.51, No.10, pp.1141-1153, Oct., 2002.

[17] S. Pack and Y. Choi, "A Study on Performance of Hierarchical Mobile IPv6 in IP-based Cellular Networks," IEICE Transactions on Communications. Vol.E87-B, No.3 pp.462-469, Mar., 2004.

[18] M. Woo, "Performance Analysis of Mobile IP Regional Registration." IEEE Transactions on Communications, Vol.E86-B,

No.2, pp.472-478, Feb., 2003.

[19] X. Zhang, J.G. Castellanos and A. T. Capbell, “P-MIP: Paging extensions for mobile IP,” ACM Mobile Networks and Applications, Vol.7, No.2, pp.127-141, Mar., 2002.

[20] Elizabet J. O'Neil, and Gerhard Weikum, “The LRU-K Page Replacement Algorithm for Database Disk Buffering,” In Proceedings of the ACM International Conference on Management of Data(SIGMOD), pp.297-306, May, 1993.

[21] J. T. Robinson and N. V. Devarakonda, “Data cache Management Using Frequency-Based Replacement,” In Proceedings of the ACM Conference(SIGMETRICS), pp.134-142, 1990.

[22] K. Cheng and Y. Kambayashi, “Advanced Replacement Policies for WWW Caching,” In Proceeding sof 1st International Conference on Web Age Information management(WAIM), June, 2000.

[23] Yong-hyun Whang, Changwoo Jung and et al., “Web-Alchemist: A Web Transcoding System for Mobile Web-Access in Handheld Devices,” Proc. SPIE Vol.4534, pp.37-47, 2001.

박 원 길

e-mail : lizzang@kict.re.kr

1996년 전북대학교 교육대학원 전산교육학과(석사)

2007년 숭실대학교 대학원 컴퓨터학과(박사)

2008년~현 재 한국건설기술연구원 박사후 연구원

관심분야 : Mobile IP, 이동통신, 멀티미디어 통신

강 의 선

e-mail : iami86@ssu.ac.kr

2002년 숭실대학교 대학원 컴퓨터학과(석사)

2007년 숭실대학교 대학원 미디어학과(박사)

2007년~현 재 숭실대학교 교양특성화대학 전임강사

관심분야 : 멀티미디어 기술, 무선 인터넷