WIRELESS COMMUNICATIONS AND MOBILE COMPUTING
Wirel. Commun. Mob. Comput. (2010)
Published online in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/wcm.1032

RESEARCH ARTICLE

A network selection scheme for multicast applications in wireless network convergence

Jaecheol Kim[1], Ji Hoon Lee[2], Hojin Lee[2], Taekyoung Kwon[2]* and Yanghee Choi[2]

[1] Department of Electronics and Computer Science, Korea Air Force Academy, Chungbuk, R.O. Korea
[2] School of Computer Science and Engineering, Seoul National University, Seoul, R.O. Korea

ABSTRACT

In wireless network convergence, each mobile host is expected to have multiple kinds of wireless interfaces. Multicast-based applications are expected to be widely deployed. In this paper, a new network selection scheme is proposed for a mobile host to select the most appropriate wireless access network to maximize user satisfaction and ISP's profit, simultaneously. We have devised a metric to measure a user's satisfaction and we also developed a 'normalized network resource' metric for system profit measure. We have compared our scheme with three other reference schemes, through simulations. Depending on the network deployment situations, our scheme exhibits one-fifth service disruption time of other reference schemes, while the resource consumption of our scheme is comparable to that of the minimum resource scheme. Overall, the gain of our scheme becomes higher as users move faster and/or the population density increases. Copyright © 2010 John Wiley & Sons, Ltd.

KEYWORDS
network selection; multicast; user satisfaction; normalized network resource

***Correspondence**
Taekyoung Kwon, School of Computer Science and Engineering, Seoul National University, R.O. Korea.
E-mail: tkkwon@mmlab.snu.ac.kr

1. INTRODUCTION AND MOTIVATION

In future wireless networks, live streaming applications are expected to be one of the killer applications. Cisco forecasts that Internet video streaming and downloads are beginning to take a larger share of bandwidth, and will grow to over 60 per cent of all consumer Internet traffic in 2013 and almost 64 per cent of the world's mobile data traffic will be video by 2013 [1,2]. User created contents (UCCs), which are gaining tremendous popularity are already utilizing live streaming services worldwide. IPTV (Internet Protocol Television) is being deployed over the Internet. Even *via* various mobile devices, many kinds of multimedia streaming contents are available [3,4]. For this kind of adaptive applications, IP multicast provides the most efficient way to utilize network bandwidth. As a result, recently, IP multicast has been revisited [3]. Other examples of applications which need the support of multicast are video conferencing, file distribution, software update and multiplayer on-line games. Especially, adaptive traffic such as live streaming applications can gain greatly from IP multicast in terms of network resources.

Another specific feature of the future wireless networks is the coexistence of heterogeneous radio access networks: WLAN, WMAN, cellular and so on. Hence, mobile hosts will have multiple radio interface cards or a single interface card which can support multiple radio technologies through software-defined radio to utilize cost efficient and quality network services [3]. In some regions, there might be just one radio access network available, and in other places, multiple access networks will be available at the same time. In the latter case, a big challenge emerges on how to select the most appropriate wireless access network in terms of user satisfaction and system resource efficiency. Some work on how to select radio access networks has been done [5–12].

Our work is unique from the viewpoint that it maximizes every single user's satisfaction and service provider's system profit at the same time under multicast support. In this paper, we show that the proposed access network selection scheme outperforms other mechanisms in terms of service disruption time. In terms of the system profit, our scheme shows near optimal efficiency. Our wireless access network selection algorithm targets adaptive applications supported by IP multicasting. Adaptive traffic is not just one application that can be supported by multicast, but it shows most clearly the effectiveness of multicast and is gaining more attention as the killer application.

We have two objectives in developing our scheme for access network selection.

- First, we aim to maximize the utility of link bandwidth and to minimize handoff-related service disruption time.
- Second, we seek to minimize system resource consumption in order to maximize system profit, where 'system' means the service provider. For the convenience of comparison, we assume that a single provider operates multiple wireless access networks.

The remainder of this paper is structured as follows. Section 2 reviews related work. Section 3 describes the metrics that are the criteria of our scheme. Section 4 presents the details of our scheme including access network selection algorithm and compared schemes. Section 5 describes the simulation setup, results and discussions. Section 6 summarizes and concludes the paper.

2. RELATED WORK

Wireless access network selection mechanisms have been examined from many viewpoints and criteria [5–12]. Heterogeneous wireless networks for the convergence era will impose many challenging issues: seamless connectivity, complex resource allocation, interference, power consumption, and security [13,14]. Each of the above issues can be boiled down to a network selection criterion. The network selection schemes in the literature can be partitioned into two categories depending on who makes the selection decision: user centric or network centric. Both categories are partially relevant to our objectives, but not applicable directly to the multicasting enabled environment, because almost all of them do not utilize multicasting. To the best of our knowledge, our work is unique in that it achieves the system resource minimization and every single user's satisfaction maximization, at the same time.

2.1. User centric approach

A user satisfaction metric is defined and a network selection algorithm considering user mobility is proposed in a heterogeneous wireless network environment in Reference [5]. This work has the objective of maximizing user satisfaction. User satisfaction function was devised by adopting bandwidth utility and satisfactory degradation factor induced by handoffs. By numerical modeling, an efficient access network selection algorithm is proposed under the assumption that two access networks are overlapped. This work, however, is irrelevant to multicast and is just focused on user satisfaction in that it has not considered overall system cost or resource consumption.

Network selection decision for non-real-time data applications is considered in Reference [9]. Here, users choose the radio access network which meets their data transfer terms best. Naturally, each user wants timely quality data delivery at a low cost. In the radio environment, data rates can never be guaranteed due to the unreliable nature of the radio links. Access network selection algorithm of this scheme is based on the prediction of the data rate of each available network. This scheme allows the user to select the network that will maximize the consumer's bandwidth requirement at a fixed price for non-real-time data, while taking into account the delays. This work is user-centric so it does not take into account the ISP's resource consumption and profit.

A novel multi-criteria network selection algorithm for always best-connected service provisioning is proposed in Reference [10]. It relies on a suitably defined cost function, which at the same time takes into account metrics reflecting both objectives of network and user preference. The strong point of this proposal is the implementation of the selection algorithm at a middleware layer; this hides both network cost computation and 4G scenario complexity from user and application layers. The effective performance observed is mainly due to the possibility of associating a weight to each cost parameter that is dynamically adapted to user preferences and profile, not only on a per-session basis, but also within the same session. This scheme is still targeted for unicasting infrastructure.

All of the above mechanisms are designed for unicasting environments, and they cannot be applied to multicast networks.

2.2. Network centric approach

Bandwidth-efficient multicasting in heterogeneous wireless networks is examined in Reference [6]. It gives the insight to utilize system resource in a cost effective way in multicast enabled communication environments. The mechanism supports the dynamic group membership and offers mobility of group members. The authors formulate the mechanism as the problem of selecting the cell and the wireless technology for each group member to minimize the total bandwidth cost of the shortest path tree. However, it is solely focused on the system's total bandwidth cost without considering individual user's satisfaction. In the future Internet, from a user's standpoint, system resource optimization is not the primary matter of concern. He or she will just be satisfied when the contents are serviced with proper QoS. Their mobility concern is also restricted to human walking speed, not covering the high speeds of vehicles.

The bandwidth requirement of services and system profit is considered at the same time for the access network selection process in Reference [7]. The authors assume multicast enabled heterogeneous wireless networks. Their mechanism is based on simple heuristics. Thus, it still lacks the element of user mobility which is the key support for future wireless networks. Therefore, these schemes can hardly be adopted for high-speed mobile environments.

In Reference [8], the authors address the access network selection algorithm that satisfies the bandwidth requirement of services, while maximizing the system profit obtained in the combined network, which is similar to the approach of Reference [7]. A heterogeneous network comprising multicast broadcast multimedia service (MBMS) of the third generation mobile terrestrial network and the digital video broadcasting transmission system for handheld terminals (DVB-H) is adopted in the study. Both networks cooperate and complement each other to improve the resource usage and to support 'one-to-many' services with their multicast and broadcast transmission capabilities. Based on this architecture, an algorithm framework is defined to solve the network selection problem for the 'one-to-many' services. Their user preference is just one parameter of bandwidth requirements, and the other important factor of handoff delay is not considered.

A dynamic network selection mechanism is presented in Reference [11], where multiple attributes and metrics are considered. Especially, the authors use kernel regression functions for the network selection under the circumstance of dynamic preference changes. Here, three kinds of utility functions are considered; availability utility, cost utility, and quality utility. Current utility and expected future utilities of other networks are compared and then the network decides whether to execute vertical handoff or not. Comprehensive metrics and utilities are considered in this scheme. However, this scheme also targets general access network selection and is not directly applicable for multicast enabled environments.

Under all-IP pervasive networking environment, a novel model to handle the network selection issue is proposed in Reference [12]. A traditional way to select a target network which is only based on the received signal strength (RSS) is not effective enough to make the best algorithm. The traffic characteristics, the user preference, and the network conditions should all be considered to maximize consumer satisfaction. Though some existing schemes do consider multiple criteria (e.g. QoS, security, connection cost, etc.) for network selection, there are still several problems unsolved. In this study, the authors obtain the necessary information of neighboring networks *via* IEEE 802.21 MIHF [15], and classify the information into two categories; then they use the non-compensatory information as a trigger for checking the compensatory information; at last, taking the compensatory information as input, they propose a hybrid analytic network process (ANP) model to rank the candidate networks. They also provide a comprehensive way to select the optimal network.

3. PERFORMANCE METRICS

As mentioned in Section 1, we have two measurable objectives; user satisfaction and system resource consumption. We seek to achieve higher user satisfaction and lower system resource consumption.

3.1. User satisfaction

User satisfactory level is directly impacted by two factors; available bandwidth (or bit rate) and handoff delay. Hence, we present two performance metric functions using those factors, respectively and then we combine them into a single satisfaction function.

3.1.1. Bandwidth utility.

The degree of user satisfaction of bandwidth requirement is defined as 'bandwidth utility'. Breslau and Shenker show that bandwidth utility can be different according to the type of applications [16]. They divide applications into three categories: rigid, elastic, and adaptive.

Rigid applications require their data to arrive within a given delay bound. Traditional telephony is the example of such applications. For these applications needing '$\underline{b}$' units of bandwidth, the bandwidth utility of given bandwidth 'b' is shown in Equation (1).

$$U(b) = 0 \text{ for all } b<\underline{b} \text{ and } U(b) = 1 \text{for all } b>\underline{b} \quad (1)$$

At the opposite end of spectrum are elastic applications. Traditional data applications such as FTP and email are the main examples. They are relatively insensitive to individual packet delays, and typically do not have hard real-time constraints. For this kind of application, though giving additional bandwidth certainly aids performance, the marginal improvement for additional bandwidth decreases in b.

The third type of application is an adaptive one such as video and audio streaming data. Our scheme targets this kind of applications to utilize the efficiency of multicasting. The bandwidth utility for adaptive applications is given as Equation (2), where K is a constant (0.62086) [16,17]. In Reference [16], the value of K is calculated to allow flows to utilize the full capacity of a link and is generally applicable to adaptive traffic.

Traditionally video and audio applications have been designed with hard real-time requirements, but most of them are adaptive for the occasional delay-bound violations and dropped packets. Even though these adaptive applications have an intrinsic bandwidth requirement, the performance degradation due to the smaller bandwidth than the intrinsic requirement is not so sharp as the case of rigid applications. From these conjectures and *Gedanken* experiment [16], Breslau and Shenker built the Equation (2) [16,17].

$$U(b) = 1 - e^{-\frac{b^2}{K+b}} \quad (2)$$

Bandwidth utility increases as the available bandwidth for a user increases, as in Figure 1. When the available bandwidth exceeds the required QoS (upper right part of the curve), the increasing ratio of satisfaction subsides. To

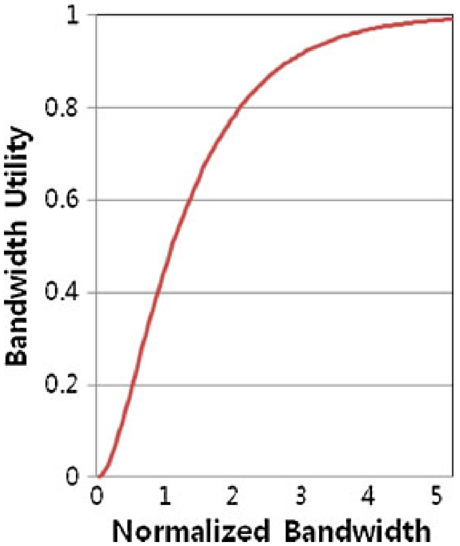


Figure 1. Bandwidth utility of adaptive applications.

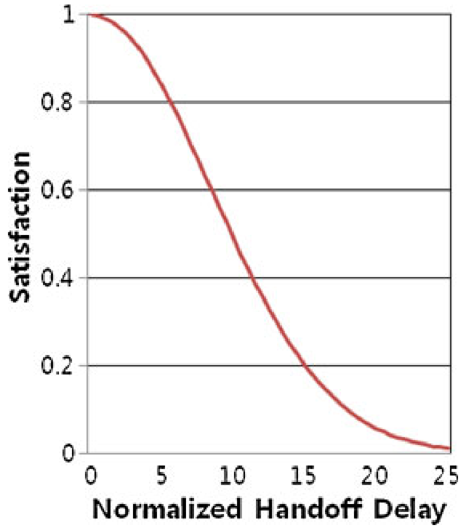


Figure 2. Satisfaction function of handoff delay.

apply Equation (2), the value of the provided bandwidth should be normalized to fit in the equation. The required bandwidth is normalized to the inflection point, where the curve changes from convex to concave in Figure 1.

3.1.2. Handoff delay.

Another factor of user satisfaction depends on the handoff delay caused by user mobility. Satisfactory level decreases as handoff delay, t_h increases. Service degradation function shows the level of satisfaction from 1 to 0 as the handoff delay increases. This satisfaction function of the handoff delay is given by Equation (3) as proposed in Reference [5]. Here, σ is a constant and its value varies according to the application characteristics; larger for adaptive ($\sigma = 5$) and elastic applications ($\sigma = 10$) and smaller ($\sigma = 2$) for real-time applications [5]. We set σ to 5 because our target applications are adaptive ones.

$$S_d(t_h) = e^{-\frac{{t_h}^2}{2\sigma^2}} \tag{3}$$

Satisfaction degradation as a function of the handoff delay is depicted in Figure 2. A user is fully satisfied when the handoff delay is 0 and so the satisfaction value is 1. At any handoff delay larger than 0, the satisfaction level is less than 1. As is the case of bandwidth utility, to apply Equation (3), we normalize the value t_h by dividing the real handoff delay by maximum tolerable delay threshold.

3.1.3. Combined satisfaction function (CSF).

We combine the above two functions into a single utility to quantify user satisfaction in one framework which we will use as one of the criteria for access network selection. The combined satisfaction function is shown in Equation (4).

$$\mathrm{CSF}(b, t_V, t_{iH}) = \left\{ \begin{array}{c} U(b)\cdots\cdots\cdots\text{no\_handoff} \\ U(b) \times \prod_{i=1}^{n} S_d(t_{iH})\cdots\cdots\text{horizontal\_handoff} \\ U(b) \times S_d(t_V) \times \prod_{i=1}^{n-1} S_d(t_{iH})\cdots\text{vertical\_handoff} \end{array} \right\} \tag{4}$$

The CSF is basically formed by multiplying bandwidth utility and service degradation function. If $U(b)$ is 0, then CSF goes to 0 even though the satisfaction function of handoff delay gives out a high value. This is the case when bandwidth utility cannot be satisfied at all. One can reason that the bandwidth utility or the satisfaction function of handoff delay should be weighted according to the strategy of an ISP or the traffic type. We leave this kind of function tuning for future work.

When a user stays in a cell and there is no handoff, a user's satisfactory level, CSF is just determined by bandwidth utility function, $U(b)$. This is the normal service mode of no handoff and is denoted by the first line of Equation (4).

The second and third lines of Equation (4) exhibit the satisfaction of handoff cases, where t_{iH} is the ith horizontal handoff delay and t_V is the vertical handoff delay. A horizontal handoff happens when a user moves from a cell to another one while keeping the access network connectivity not changed. On the other hand, vertical handoff occurs when the access network connectivity is changed between different networks. Here, CSF is derived by the multiplication of a bandwidth utility function and anticipated service degradation caused by handoffs throughout the entire session, thereafter. The term 'n' is calculated by session duration, network cell coverage and moving speed. Intuitively, '*n*' represents the number of total horizontal handoffs throughout the session duration while keeping the average moving speed. In our simulation, session duration is set to 30 min of simulation run time but in reality the session duration needs to be pre-databased from statistical analysis of session type and user behavior. Especially, once a vertical handoff occurs, handoffs thereafter will be horizontal ones, as assumed in the third line of Equation (4). Beware that the third line does not mean vertical handoff actually occurs just one time. Equation (4) is used for pre-estimation and comparison. In other words, the comparison is for the sake of deciding whether changing the access network and keeping it throughout the session will give more user satisfaction or not. At every handoff occasion, the second line and third line of Equation (4) are pre-estimated, and used as the input value for the access network selection algorithm, which is described in the Section 4.1.

3.2. System resource

One of the reasons why IP multicast is not yet deployed so much is that there are no significant incentives for ISPs. For successful deployment of multicast, there should be a measurable profit gain for ISPs. In a multicast enabled environment of heterogeneous wireless access networks, adding a user to the same content of an access network requires no additional cost, because that content is broadcast on the air. As a result, simple bandwidth summation of every single user does not exactly reflect the system resource. From the ISP's standpoint, if it can provide the same content to more users by reduced total bandwidth consumption without sacrificing quality, its profit gain will increase. Here, we assume that a single ISP has the control over heterogeneous wireless access networks as stated in Section 1. The objective of our system resource metric is to select the network which minimizes that metric for an ISP's profit gain because system resource metric and system profit are inversely proportional to each other.

In heterogeneous wireless networks environment, each wireless technology has unique cell coverage coupled with bandwidth and some other factors, which are relevant to costs. Therefore, for the comparison, we developed a 'normalized network resource' metric, NNR.

$$\mathrm{NNR} = \frac{\text{Session Bandwidth} \times \text{Cell Coverage}}{\text{Network Capacity} \times \text{Number of Users}} \quad (5)$$

In Equation (5), 'Session Bandwidth' is the required bandwidth for a multicast session, and 'Network Capacities' for WLAN, WMAN and the cellular network are 11, 3, and 1 Mbps, respectively, and are subject to change according to the evolution of technologies. 'Cell Coverage' is the area of one cell of each network and this coverage is tightly coupled with power consumption and the cost for building up infrastructures. It should be normalized by the 'Number of Users' for cost computation because the increased number of current users means increased ISP's profit. We use Equation (5) as the network resource comparison metric.

4. PROPOSED SCHEME

4.1. Network selection algorithm for multicast sessions

The proposed algorithm is network centric in that it utilizes full access network information such as the entire network map, bandwidth utility, the number of current users for a multicast session, pre-estimation of anticipated handoff delays and so on. We combine CSF and NNR into a single metric for the sake of network selection; that is, the most preferable network has the highest CSF/NNR. The node which executes the selection algorithm is a dedicated selection server of the ISP in charge of the access networks.

The proposed algorithm has the following procedure. When a call request (or handoff request) arrives, if there is no multicast session, the algorithm selects the access network of maximum CSF/NNR. If there is only one access network that services the requested session, then the CSF of that access network is evaluated and if it is greater than the lower bound, that network is selected. However, if CSF of the network is less than lower bound, other access network of maximum CSF/NNR is selected. Each wireless access network has its own maximum physical connection capacity. Actually, CSF_lower_bound is set to 0 for the case when a network cannot accommodate a new user physically and thus bandwidth utility becomes 0. In usual case, CSF is larger than 0, but when the maximum number of users for a cell of a network exceeds the threshold, CSF goes to 0. If there are multiple access networks that service the requested session, then the algorithm selects the network of maximum CSF/NNR among the networks in service. The terms CSF and NNR can be combined by some weights according to the ISP's policy. Figure 3 depicts the algorithm in a procedural flow chart.

(1) When a call request (including user preferences (Min BW, moving speed)) arrives, or a handoff is going to occur, list the available access networks with 2 attributes of (Combined Satisfaction Function (CSF), Normalized Network Resource (NNR))
(2) if (there's no multicast session)
(3) select the access network (AN) for maximum (CSF/NNR)
(4) else if (there's already a multicast session only on one AN)
(5) if (CSF > CSF_lower_bound) select that AN
(6) else select the AN among other ANs for maximum (CSF/NNR)
(7) else if (there are multiple sessions through multiple ANs)
(8) select the AN among the above multiple ANs for maximum CSF

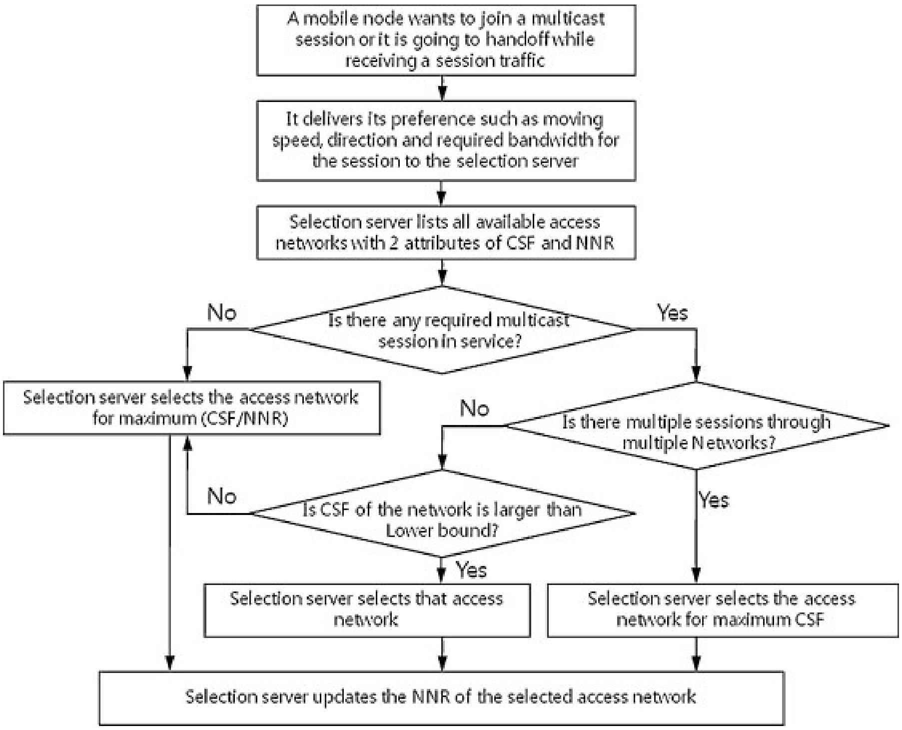


Figure 3. Access network selection procedure.

(9) end if
(10) update NNR of the selected network.

A call request includes some user preferences such as minimum bandwidth requirement and moving speed of the user for the ISP to select the most appropriate network. Nowadays, most smart phones are equipped with GPS modules and usually have even navigation applications. For such mobile devices, the estimation of moving speed and direction is not that burdensome. But, for the devices without GPS module, moving pattern estimation requires expensive energy consumption. In this case, effective mobility prediction algorithms of previous research can be used considering the energy-constrained characteristics of mobile devices [18,19]. These preferences are assumed to be conveyed to the ISP by some protocols, such as IP multimedia subsystem (IMS) or other protocols, which is out of the scope of this paper.

4.2. Reference schemes

We devise three other schemes to be compared with ours.

The *MaxUtility* scheme is designed to select an access network which provides the largest available bandwidth, which has the same approach as the work of Reference [9]. It maximizes every single user's bandwidth utility and does not consider the handoff delay effects. At any handoff, it usually prefers WLAN to WMAN and WMAN to the cellular network. Recall that the nominal bandwidth capacities of WLAN, WMAN and the cellular network are 11, 3, and 1 Mbps, respectively. *MaxUtility* is a user centric approach and does not consider existing users of the requested session. As a result, it does not fully utilize the efficiency of multicast.

The *MaxUtility-SD* scheme considers service degradation caused by handoffs together with each user's bandwidth utility. While this scheme also uses the CSF of our proposed scheme for service degradation estimation, it does not consider NNR or existing sessions. Because neither *MaxUtility* nor *MaxUtility-SD* considers the users in the existing sessions, they inevitably lack the efficiency of multicast.

The *MinResource* scheme selects an access network which minimizes NNR and is a variant of the scheme presented in Reference [6]. This scheme is network centric in that it is not concerned with each user's satisfaction but is focused on overall system resource efficiency and best utilizes the efficiency of multicast. Therefore, in terms of system resource consumption, this scheme is optimal.

5. SIMULATIONS

5.1. Simulation environments

The simulation environment is constructed by C language and the heterogeneous wireless access networks for our simulations are composed of WLAN, WMAN (Mobile

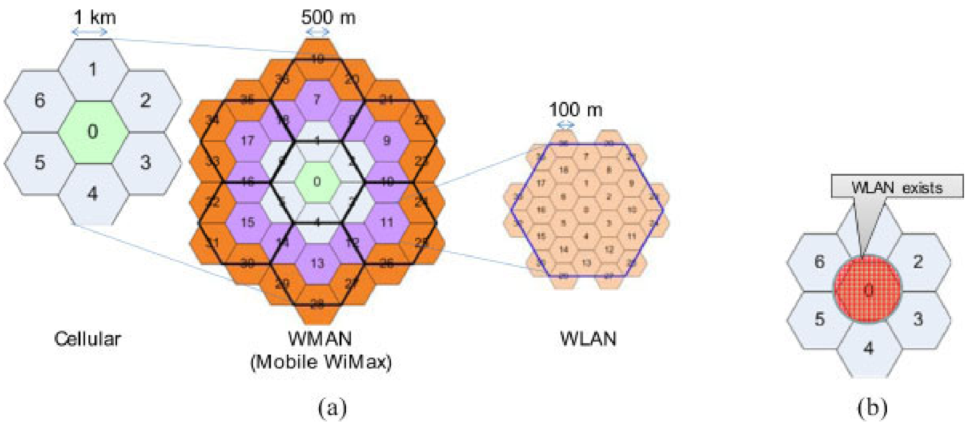


Figure 4. Network models: (a) full coverage model; (b) hot spot model.

WiMAX) and the cellular network. We simulate a multicast service area comprising of 7 hexagonal cellular-network cells, which is overlapped by 37 WMAN cells. The service area is actually wrapping-around to remove the boundary effect. For WLAN, we set up two types of network coverage models: full coverage model and hot spot model. The former assumes that the entire service area is also covered by 700 WLAN cells so that the area is fully overlapped as depicted in Figure 4a. Although this model may be unrealistic in the sense that WLAN covers the entire area, it is useful for figuring out how our access network selection mechanism works. The latter is a hot spot model where WLAN covers only the center of the service area, namely the hot spot one cellular cell which is depicted in Figure 4b. Only one cell at the center is covered by 100 WLAN cells. The radii of a single cell of WLAN, WMAN and the cellular network are set to 100 m, 500 m and 1 km, respectively.

There are six simultaneous multicast sessions that a user may join, and each session is transmitted at a rate of 500 kbps.† Our simulation uses a random waypoint model, where each user moves along a straight line from one waypoint to the next for 100 s; the next waypoint is randomly chosen between −90 degree and +90 degree of the current direction. In the meantime, the user speed changes every 50 s, which is uniformly distributed. The simulation duration of each run is 30 min.

To accommodate the various settings of user distribution and mobility, we establish four different user profiles‡ that vary with the average speed and the number of users. Profile A stands for the users of walking speed mobility receiving an unpopular session (sparse population) and profile B represents the case of same mobility and a popular session (dense population). Profiles C and D are set up for high speed mobility (vehicular speed). Table I summarizes the profile setup and the input simulation parameters.

We investigate two metrics as the evaluation of the network selection algorithm performance. One is the average service disruption time of users and the other is the cumulative NNR. For each user, the sum of all horizontal and

Table I. Simulation parameters.

Cellular Network	Size (number of cells)	7
	Cell radius	1 km
	Capacity	1 Mbps/cell
	Handoff delay	100 ms
WMAN	Size (number of cells)	37
	Cell radius	500 m
	Capacity	3 Mbps/cell
	Handoff delay	150 ms
WLAN	Size (number of cells)	925
	Cell radius	100 m
	Capacity	11 Mbps/cell
	Handoff delay	300 ms
Vertical handoff delay	WLAN ↔ Cellular	500 ms
	WMAN ↔ Cellular	500 ms
	WLAN ↔ WMAN	500 ms
Number of users	Profile A	10
	Profile B	1000
	Profile C	10
	Profile D	1000
	Others	10 ∼ 2000
Average speed	Profile A	6 km/h
	Profile B	6 km/h
	Profile C	120 km/h
	Profile D	120 km/h
	Others	0 ∼ 120 km/h
Mobility model		Random waypoint model
Multicast source rate		0.1, 0.5, 1, or 2 Mbps
Simulation duration		30 min

† Later, we extend the experiment to include the source rate of 0.1, 0.5, 1, and 2 Mbps in Section 5.2.2.

‡ We also explore the effects of varying number of users and varying average speed in Section 5.2.2.

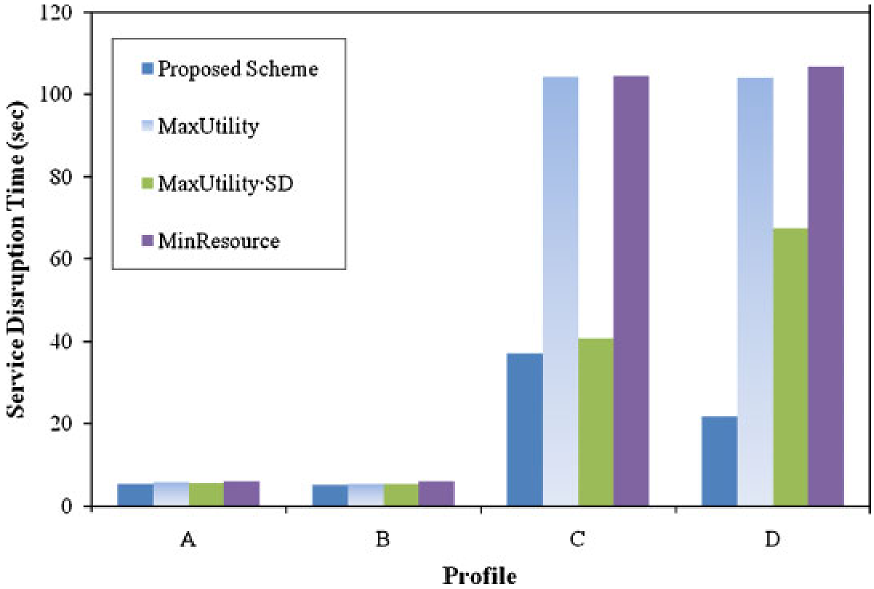


Figure 5. Service disruption time under the full coverage scenario.

vertical handoff delays experienced in the simulation will be the service disruption time. To quantify handoff delays, various materials of previous research are referred [20–27]. The work by Tripathi *et al*. presents two kinds of handoffs in cellular networks; network-controlled handoff and mobile-controlled handoff whose delays vary from 100 to 200 ms [27]. The handoff delay requirement for WMAN (Mobile WiMAX) of 2.3 GHz portable wireless broadband access is standardized as 150 ms [22]. An empirical measurement of link layer handoff delays for WLAN is presented in Reference [25]. They provide six kinds of experiment configurations of various APs and NICs from different vendors. Average handoff delays vary from 58 to 396 ms depending on the setups. From the experiment we chose the conservative value of 300 ms handoff delay for our simulation. The vertical handoff delays between WLAN and cellular networks are estimated from 480 to 600 ms [20,23,26]. The work by D. Kim *et al*. presents the vertical handoff delay between WMAN and WLAN as 610 ms [21]. To the best of our knowledge, the vertical handoff delay between WMAN and cellular networks has not been published, but the desired handoff delay is presented as 300 ms in Reference [24]. Actual vertical handoff delays between access networks are subject to change according to the advances in technologies, and we set the approximated averaged value as 500 ms for the simulations. The cumulative NNR is calculated by adding the NNR value measured per second.

5.2. Simulation results

We analyzed the performance of our scheme in terms of the service disruption time and NNR, each of which measures a mobile user's satisfaction and the ISP's profit, respectively. The three other schemes compared to ours are *MaxUtility, MaxUtility-SD*, and *MinResource*. The simulations are executed under two scenarios; a full coverage model and a hot spot model. Both models are incorporated with four user profiles.

5.2.1. Full coverage model.

The Full coverage scenario represents each of the wireless access networks covering the entire service area. Service disruption time is compared in the Figure 5. Depicted service disruption time is the averaged value of disruption times which every single user experiences during the simulation (30 min).

Profiles A and B show no meaningful differences among the selection schemes. Under low speed mobility, all the schemes, including ours, select WLAN all the time, because its bandwidth utility is high and handoffs themselves are limited. The NNR of WLAN is also significantly smaller than those of other access networks.

With high speed mobility, the number of handoffs increases. Table II shows the residence time ratio of each access network for each scheme. We cumulate each user's residence time in each access networks during the simulation time (30 min) and represent them in percentage. To reduce handoff delay, the proposed algorithm selects the cellular network compared to the other schemes. Under the profile C of sparse population, the residence duration of the cellular network is 1%. However, under the profile D of the same speed and dense population, the cellular network is selected more and the residence duration increases from 1 to 25%. The service disruption time of our proposed scheme under the profile D is 14.1% of *MaxUtility,* 21.8% of *MaxUtility-SD*, and 13.8% of *MinResource*. It can be

Table II. Residence time ratio of each access network by schemes.

Profiles	A				B				C				D			
	6 km/h				6 km/h				120 km/h				120 km/h			
	10 users				1000 users				10 users				1000 users			
Scheme	P	M	MS	MR	P	M	MS	MR	P	M	MS	MR	P	M	MS	MR
Cellular	0	0	0	0	0	0	0	1	1	0	1	0	25	0	4	1
WMAN	0	0	0	0	0	0	0	1	78	0	75	0	72	0	42	1
WLAN	100	100	100	100	100	100	100	98	21	100	24	100	3	100	54	98

P: *Proposed Scheme*
M: *MaxUtility*
MS: *MaxUtility-SD*
MR: *MinResource*

said that comparative satisfaction level is higher as the users move faster and the population is denser (profiles C and D).

MaxUtility and *MinResource* show little changes in network selection tendency without regard to the moving speeds because they do not consider handoff delays. In case of the sparse population (profiles A and C), our proposed scheme does not show large difference from *MaxUtility-SD*. Under sparse population, our scheme is limited to utilize the efficiency of multicasting, which gives little room for user aggregation.

The NNR consumption of each scheme is shown in Figure 6. As expected, under the profiles A and B, this term also shows no meaningful differences. But, in the case of the profiles C and D of high-speed mobility, the proposed scheme consumes much more NNR than other schemes. This result comes from the fact that the other schemes select mainly WLAN, because its NNR is significantly smaller than the cellular network or WMAN. As stated before, our proposed scheme is doubly targeted for CSF maximization and NNR minimization and so under the profiles C and D, it selects the cellular network and WMAN more often than other schemes to reduce service disruption time caused by handoffs. From Table II, we find that for the users who move with high speed, the proposed scheme selects the cellular and WMAN dynamically to reduce the number of handoffs and handoff delays.

5.2.2. Hot spot model.

Figure 7 shows the total service disruption time under the hot spot scenario. Our scheme shows the shortest handoff delay under any profile. Especially under the profile D, our scheme substantially outperforms the other schemes. Table III shows the comparative service disruption time of the proposed scheme. Under low-speed mobility (profiles A and B), our scheme exhibits shorter service disruption time compared with that of the other schemes. Just like the full coverage model, in case of the sparse population profile C, our proposed scheme does not show large difference from MaxUtility-SD. Under sparse population, our scheme is limited to utilize the efficiency of multicasting, which

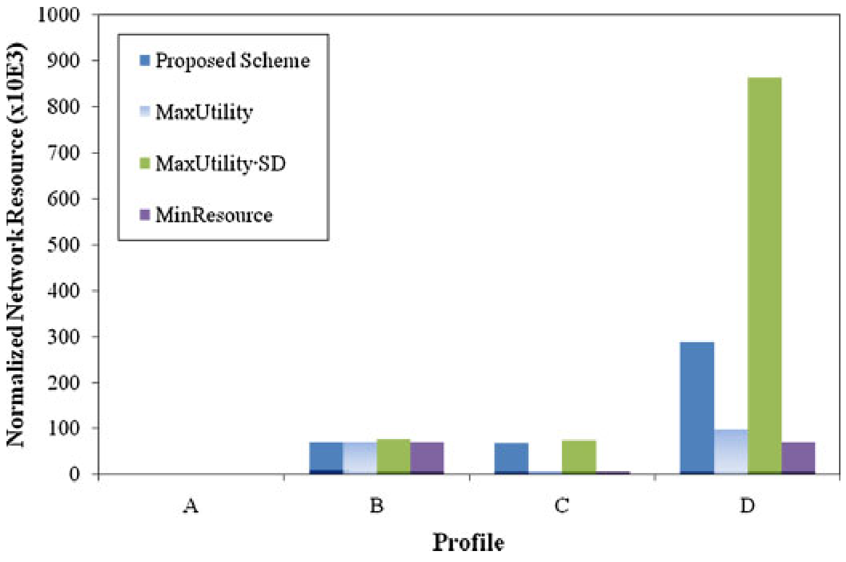


Figure 6. NNR under the full coverage scenario.

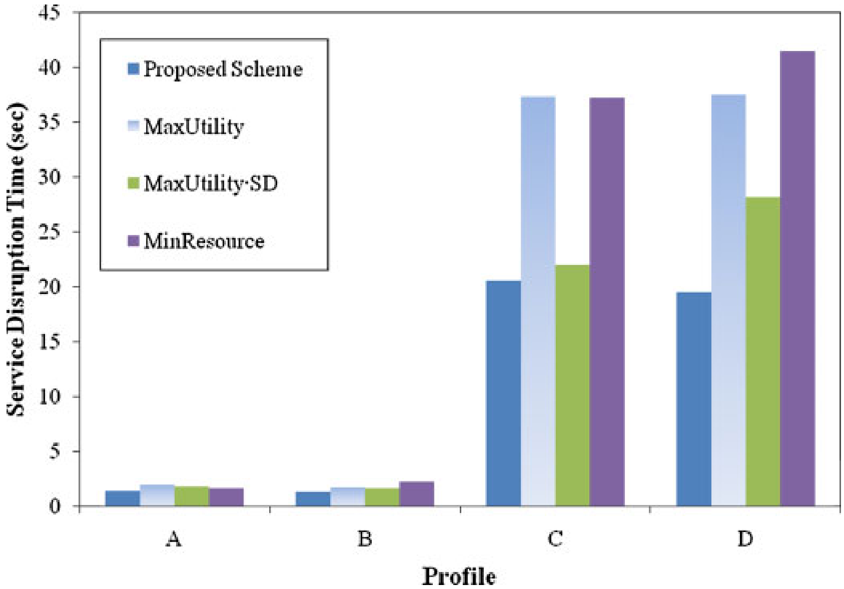


Figure 7. Service disruption time under the hot spot scenario.

gives little room for user aggregation. However, in dense populations and high speed mobility (profile D), it presents much shorter service disruption time. The results show that the key criterion of our algorithm, CSF/NNR is dynamically influenced by the number of users and the moving speed to consider the service disruption.

Figure 8 shows the NNR consumption of the four schemes. For all cases, our scheme significantly outperforms *MaxUtility* and *MaxUtility-SD*. Since both *MaxUtility* and *MaxUtility-SD* consider the bandwidth utility of each user, they consume more network resource than our scheme and *MinResource*. In particular, our scheme highlights its performance gain in the profile D, which has dense user distribution with high speed mobility. Obviously, *MinResource* is the best, since it minimizes the total network usage. However, our scheme exhibits near optimal efficiency of the *MinResource* scheme. Table IV shows the surpassing NNR ratio of our proposed scheme when compared with the *MinResource* scheme, which is optimal in terms of NNR.

In conclusion, our scheme presents best user satisfaction in terms of service disruption time and exhibits near optimal system resource usage. Comparative satisfaction level grows even higher when the moving speed is high (profiles C and D).

5.2.3. User speed, number of users and source rate.

In addition to the basic user profiles, we extend the simulation to find out the effects of some factors: user speed, number of users, and source rate. Firstly, we examine the influence of the user speed changes. We just focus on user mobility while fixing other factors; source rate is 0.5 Mbps, network configuration is full coverage model, and the number of users is 1000. Figure 9 shows the result of user speed change. Here we observe that as the user mobility speeds up, the performance gain of the proposed scheme also increases. As the user speed increases, the number of handoffs also increases. In this high-speed case, the criterion CSF is more impacted by the handoff delay. Thus, the CSF can be increased by selecting the cellular network more, which leads to reduced handoffs.

Secondly, we observed the influence of the number of users. Here, we fix other factors except that we set the user speed to 120 km/h. Figure 10 shows the result. As expected,

Table III. Comparative service disruption time of proposed scheme.

Profiles	A			B			C			D		
	6 km/h			6 km/h			120 km/h			120 km/h		
	10 users			1000 users			10 users			1000 users		
Scheme	M	MS	MR	M	MS	MR	M	MS	MR	M	MS	MR
Ratio of Service disruption time (%)	69	74	83	77	79	58	55	94	55	52	69	47

M: *MaxUtility*
MS: *MaxUtility-SD*
MR: *MinResource*

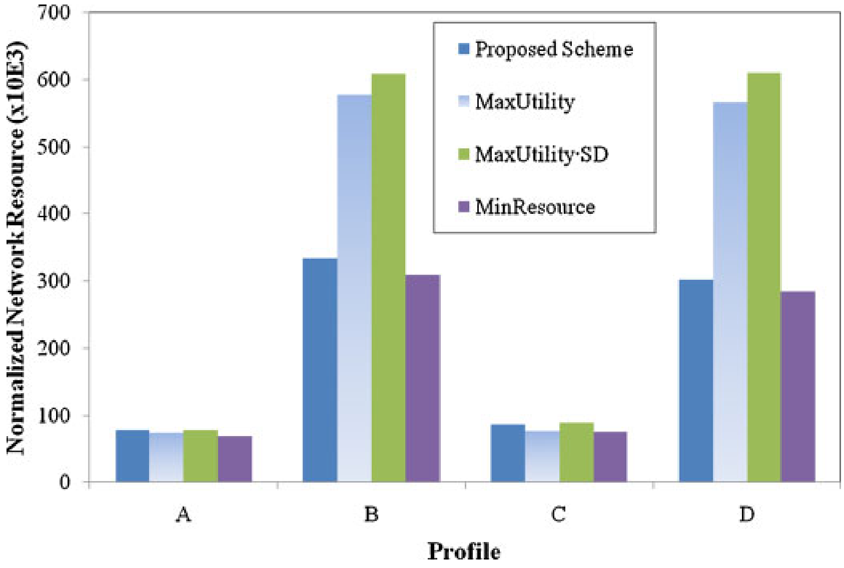


Figure 8. NNR under the hot spot scenario.

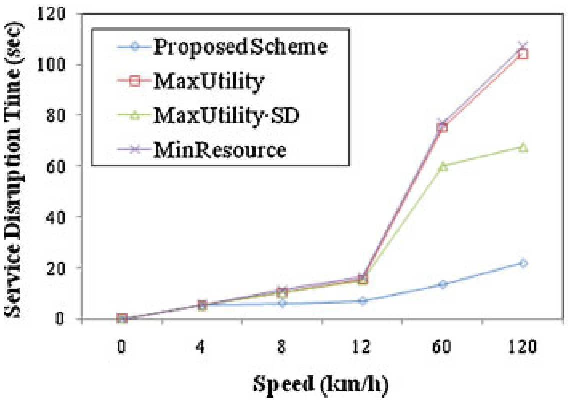
Table IV. NNR optimality of the proposed scheme.

Profiles	A	B	C	D
Proposed scheme	77.661	333.594	85.791	302.899
MinResource (Optimal)	69.116	309.509	75.022	284.400
Excess ratio over optimal consumption	12.4%	7.8%	14.4%	6.5%

both of service disruption time and NNR of the proposed scheme show better performance as the number of users grows. Very special feature of this experiment is that not just the comparative value but also even the absolute values of service disruption time of the proposed scheme improve as the number of users increases. The simple criterion of our access network selection, CSF/NNR can adapt dynamically from its criteria. While the number of users reaches to 500, WMAN and cellular network are selected more but total NNR itself increases. From that point, absolute value of NNR decreases due to CSF/NNR dynamism because NNR can be reduced by the division of the number of users (see Equation (5)) and CSF can be increased by decreased handoffs.

Anyway, our proposed scheme best utilizes the multicasting property and the efficiency of the scheme is highlighted when the population is dense.

Lastly, we observe the effects of the source rate change as is shown in Figure 11. Four kinds of source rate were examined: 0.1, 0.5, 1, and 2 Mbps. Here we fix other factors; user speed is 120 km/h and the number of users is 1000. Under any source rate, the proposed scheme shows the shortest service disruption time. However, as the source rate increases, the gaps between the schemes diminish in terms of both service disruption time and NNR. Especially, the NNR of the proposed scheme and *MaxUtility-SD* decreases when source rate increases from 1 to 2 Mbps. When the source rate becomes 2 Mbps, the cellular network cannot be selected at all, because its maximum capacity is 1 Mbps.



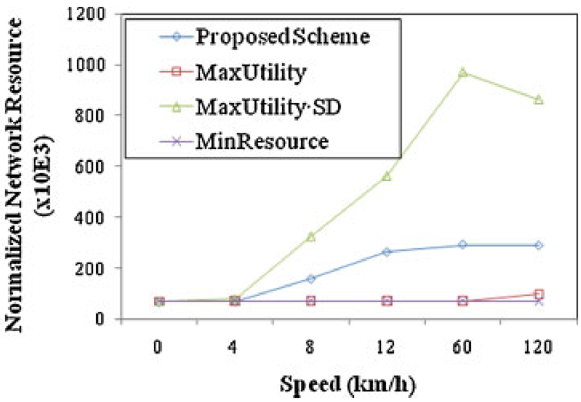


Figure 9. The effects of user speed.

DOI: 10.1002/wcm

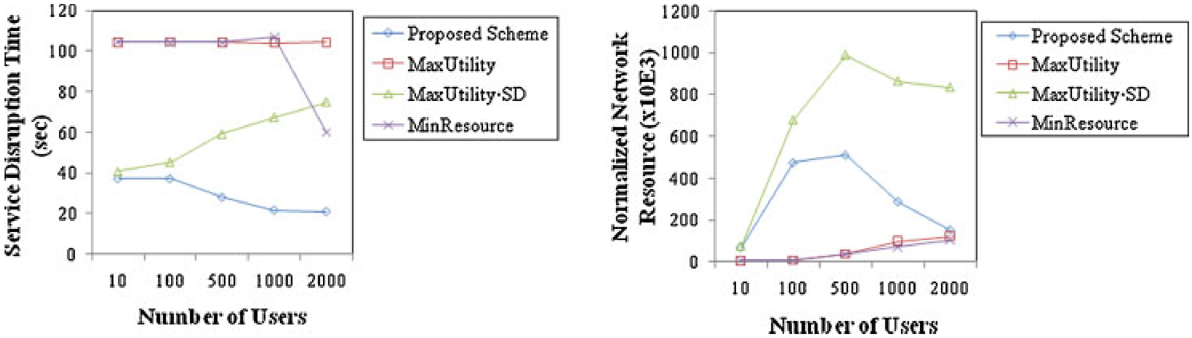


Figure 10. The effects of the number of users.

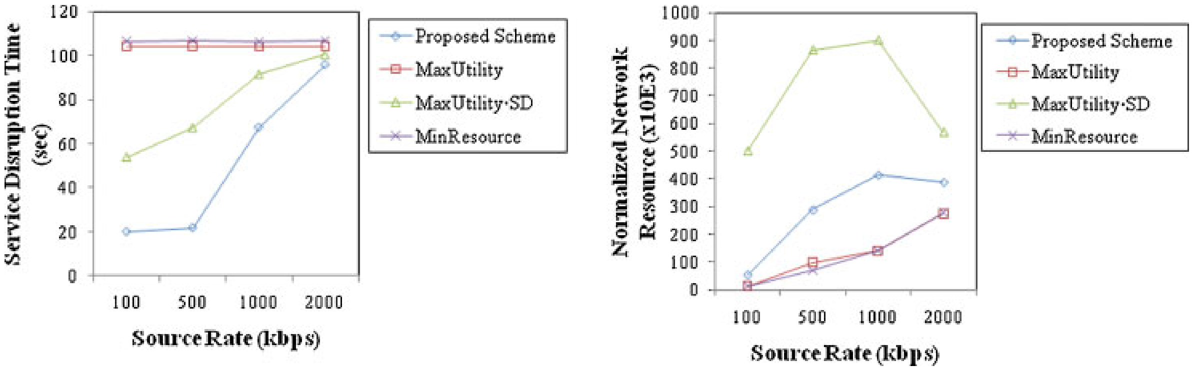


Figure 11. The effects of the source rate.

Even the WMAN can service just one session, and so five other session should be serviced through WLAN. This situation limits the maximum number of sessions possible in each network.

5.3. Discussions

Under the full coverage scenario, our proposed algorithm substantially outperforms other schemes in terms of service disruption time. NNR consumption is located in the midrange, but this can be justified by the tradeoff with the service disruption time because service disruption time has a greater effect on user satisfaction. In the more realistic environment, the hot spot scenario, our scheme shows simultaneously the shortest service disruption time and near optimal NNR consumption efficiency.

We incorporated CSF and NNR into one network selection framework, which can be utilized by ISPs for their optimal resource usage while providing users measurable satisfaction. CSF is composed of bandwidth utility and service disruption time prediction. NNR is impacted dynamically by the number of users and the moving speed of them, which results in higher satisfaction in the case of dense population. Service disruption time of our proposed scheme also shows higher performance, when the moving speed is high and the population is dense.

But in case of source rate increase, each of service disruption time and NNR consumption converges regardless of the schemes. It is because as source rate increases, the selection choice is limited physically depending on the network capacity.

Two attributes of CSF and NNR can be finely tuned according to the ISPs' strategies. We present just the preliminary framework itself.

6. CONCLUSIONS

One of the specific features of the future Internet is the heterogeneity of wireless access networks and the multicast dominance. In this paper, we propose a scheme for the efficient wireless access network selection for the multicast services in wireless network convergence. Here, efficiency is approached from two sides; user's standpoint and ISP's standpoint. We developed metrics to measure user satisfaction and ISP's resource consumption. Combined satisfaction function (CSF) combines the bandwidth utility and the satisfaction degradation caused by service disruption into one framework. Normalized network resource (NNR) exhibits the reverse proportionality of ISP's profit by incorporating requested bandwidth, network capacity, cell coverage and the number of users in a session. Our proposed algorithm efficiently utilizes the multicast properties and the metrics we devised. Through simulations, our scheme is shown to minimize service disruption time and consumes NNR almost optimally when compared with other schemes under WLAN hot spot scenario. This work is unique in that it is designed for the multicast enabled future wireless networks and has a special application target. The result of this work is a strong incentive for the ISPs to deploy multicast in that they can gain profits by reducing resources, while providing sufficient satisfaction to every single user of their networks.

Even in the case of unicasting, the objectives of the proposed scheme are still effective. First, for every single user, the scheme tries to maximize the CSF (Combined Satisfaction Function) of bandwidth utility and handoff delay. Second, the scheme tries to minimize NNR (Normalized Network Resource) with the 'number of users' factor set as 1.

In the future, more traffic types should be incorporated into this scheme and the fine tuning of the metrics, with regards to the various business strategies of the ISPs, is also needed. The scheme can be improved much finely by considering such factors as service characteristics, packet scheduling and QoS requirement related to the user satisfactions. Nevertheless, the framework for user satisfaction maximization and system resource minimization is always effective.

ACKNOWLEDGEMENTS

This work was supported by the IT R&D program of MKE/KEIT [10035245: Study on Architecture of Future Internet to Support Mobile Environments and Network Diversity] and NAP of Korea Research Council of Fundamental Science & Technology.

REFERENCES

1. Cisco. Cisco Visual Networking Index: Forecast and Methodology, 2008–2013, White paper, June, 2009.
2. Schonfeld E. Cisco: By 2013 Video Will Be 90 Percent Of All Consumer IP Traffic And 64 Percent of Mobile, www.techcrunch.com/2009/06/09/cisco-by-2013-video-will-be-90-percent-of-all-consumer-ip-traffic-and-64-percent-of-mobile
3. Ratnasamy S, Ermolinskiy A, Shenker S. Revisiting IP multicast, *Proceedings of SIGCOMM'06*, 2006.
4. Eyal M. The Streaming Media Handbook, Prentice Hall: New Jersey, 2002.
5. Kafle VP, Kamioka E, Yamada S. Maximizing user satisfaction based on mobility in heterogeneous mobile multimedia communication networks. *IEICE Transactions on Communications* 2005; **E88-B**(7): 2709–2717.
6. Yang D, Chen M. Efficient resource allocation for wireless multicast. *IEEE Transactions on Mobile Computing* 2008; **7**(4).
7. Gluhak A, Chew KA, Moessner K, Tafazolli R. Multicast bearer selection in heterogeneous wireless networks. *Proceedings of ICC'05* 2005; 1372–1377.
8. Huang L, Chew KA, Tafazolli R. Network selection for one-to-many services in 3G-broadcasting cooperative networks. *Proceedings ofVTC '05* 2005; 2999–3003.
9. Ormond O, Perry P, Murphy J. Network selection decision in wireless heterogeneous networks, *IEEE 16th International Symposium on PIMRC 2005*, Sept. 2005.
10. Iera A, Molinaro A, Campolo C, Amadeo M. An access network selection algorithm dynamically adapted to user needs and preferences, *IEEE 17th International Symposium on PIMRC 2006*, Helsinky, Finland, Sept. 2006.
11. E van den Berg, P Gopalakrishnan, B Kim and B Lyles. Dynamic Network Selection using Kernels, *Proceedings of IEEE International Conference on Communications '07*, 2007.
12. Liu Y. Access network selection in a 4G networking environment, University of Waterloo, *Thesis of Master degree*, Jan. 2008.
13. SPG Media Ltd., 4G: Challenges and Opportunities, *Global Semi-conductor Forum*, Feb. 2007.
14. Gazis V, Alonistioti N, Merakos L. Toward a generic "always best connected" capability in integrated WLAN/UMTS cellular mobile networks. *IEEE Wireless Communications* 2005; **12**: 22–29.
15. IEEE P802.21/D14 Draft standard for local and metropolitan area networks: media independent handover services, Sept. 2008.
16. Breslau L, Shenker S. Best effort versus reservation: a simple comparative analysis. *ACM Computer Communications Review* 1998; **28**: 131–143.
17. Shenker S. Fundamental design issues for the future internet. *IEEE Journal on Selected Areas in Communications* 1995; **13**(7): 1176–1188.
18. Abu-Ghazaleh H, Alfa AS. Mobility prediction and spatial-temporal traffic estimation in wireless networks, *IEEE 67th Vehicular Technology Conference*, 2008.
19. Mehbodniya A, Chitizadeh J. User location estimation in wireless networks with application to vertical handoff using pattern recognition, *Proceedings of the 10th IEEE Symposium on Computers and Communications*, 2005.
20. Bhaskara M, Murthy R, Phiri FA. Performance analysis of downward handoff latency in a WLAN/GPRS interworking system. *Journal of Computer Science* 2005; **1**(1): 24–27.
21. Do-Hyung Kim, Won-Tae Kim, Hwan-Gu Lee, Sun-Ja Kim, A performance evaluation of vertical hanover architecture with low latency handover, *International*

Conference on Convergence and Hybrid Information Technology '08, 2008.
22. Hong D. 2.3GHz Portable Internet (WiBro) for Wireless Broadband Access, ITU-APT Regional Seminar, 2004.
23. Luo H, Jiang Z, Kim B-J, Henry P. Integrating wireless LAN and cellular data for the enterprise. *IEEE Internet Computing* 2003; **7**: 25–33.
24. Kim K, Kim Y, Kim J, Lee S. PMIP based WiBro-HSDPA seamless mobility support. *KNOM Review* 2008; **11**(1).
25. Mishra A, Shin M, Arbaugh W. An empirical analysis of the IEEE 802.11 MAC layer handoff process. *ACM SIGCOMM Computer Communication Review* 2003; **33**(2).
26. Parikh H, Chaskar H, Trossen D, Krishnamurthi G. Seamless handoff of mobile terminal from WLAN to cdma2000 network, *Procedings of IEEE 3G Wireless2003*, 2003.
27. Tripathi ND, Reed JH, VanLandingham HF. Handoff in cellular systems, *IEEE Personal Communications*, Dec. 1998.

AUTHORS' BIOGRAPHIES

Jaecheol Kim is an Assistant Professor in the Department of Electronics and Computer Science, Korea Air Force Academy where he has been teaching from 1995. He received his Ph.D., M.S., and B.S. degrees in Computer Science and Engineering from Seoul National University in 2009, 1995, and 1993, respectively. During his Ph.D. commissioned education from 2002 to 2005, he was with Multimedia and Mobile Communications Lab., School of Computer Science and Engineering, Seoul National University. His recent research interest lies in wireless technology convergence, military MANET, military data-link, the Internet traffic measurement, and mobile multicasting.

Ji Hoon Lee (jhlee@mmlab.snu.ac.kr) received a B.S. degree in industrial engineering and a B.S. degree in computer science from POSTECH in 2000. He has worked on a development team for Ethernet switching and IP routing devices for the year 2000-2003, and during 2004-2005 he was a senior engineer for developing the MAC layer of an IEEE 802.16 based WiMAX base station at POSDATA R&D center, USA. He is currently pursuing a Ph.D. degree in the School of Computer Science and Engineering at Seoul National University. His research interests are in network planning for multimedia multicasting over broadband wireless networking such as WiMAX and 3GPP LTE, and autonomous femto-cellular network organization.

Hojin Lee. He received a B.S. degree in computer science and engineering from Seoul National University in 2003. He is currently working toward a Ph.D. degree at Multimedia and Mobile Communications Lab., School of Computer Science and Engineering, Seoul National University. He was a visiting researcher at Network Modeling Lab., Simon Fraser University, Canada, and at Computer Networking and Multimedia Research Group, University of Minnesota, USA. His research interests include routing protocols in VANETs, wireless broadband access networks, and energy-efficient MAC protocols in WSNs.

Taekyoung Kwon is an Associate Professor in the Multimedia and Mobile Communications Laboratory, School of Computer Science and Engineering, Seoul National University. He received his Ph.D., M.S., and B.S. degrees in computer engineering from Seoul National University in 2000, 1995, and 1993, respectively. He was a visiting student at IBM T. J. Watson Research Center in 1998 and a visiting scholar at the University of North Texas in 1999. His recent research areas include radio resource management, wireless technology convergence, mobility management, and wireless sensor networks.

Yanghee Choi received his B.S. in Electronics Engineering from Seoul National University, M.S. in Electrical Engineering from Korea Advanced Institute of Science and Technology, and Ph.D. of Engineering in Computer Science from Ecole Nationale Superieure des Telecommunications (ENST) in Paris, in 1975, 1977, and 1984 respectively. Before joining the School of Computer Engineering, Seoul National University in 1991, he was with the Electronics and Telecommunications Research Institute (ETRI) in 1977–1991, where he served as the Director of the Data Communication Section, and the Protocol Engineering Center. He was a research student at Center National d'Etude des Telecommunications (CNET), Issy-les-Moulineaux, in 1981–1984. He was also a Visiting Scientist to IBM T. J. Watson Research Center in 1988–1989. He is now leading the Multimedia and Mobile Communications Laboratory in Seoul National University. He is president of the Korean Institute of Information Scientists and Engineers. He is also chair of the Future Internet Forum. He is regular member of the National Academy of Engineering of Korea, and Korea and Korean Academy of Science and Technology. His research interest lies in the field of Future Internet.