논문 15-40-06-12 The Journal of Korean Institute of Communications and Information Sciences '15-06 Vol.40 No.06
http://dx.doi.org/10.7840/kics.2015.40.6.1081

Performance Assessment of an Access Point for Human Data and Machine Data

Hoon Lee[•°]

ABSTRACT

This work proposes a theoretic framework for the performance assessment of an access point in the IP network that accommodates MD (Machine Data) and HD (Human Data). First, we investigate typical resource allocation methods in LTE for MD and HD. After that we carry out a Max-Min analysis about the surplus and deficiency of network resource seen from MD and HD. Finally, we evaluate the performance via numerical experiment.

Key Words : EoI, MTC, QoS, traffic fluctuation, network dimensioning, performance assessment

Ⅰ. Introduction

Recently, the ecosystem of the Internet is evolving toward EoI (Everything over Internet) where various kinds of data such as human data and machine data coexist over the same network.

One of the most important issues in the current Internet is increasing volume of traffic and uncertainty in traffic variation, which is originated from the video service and introduction of MTC (Machine Type Communication) in addition to the conventional HTC (Human Type Communication).

The increasing volume of traffic can be easily coped with the deployment of wider bandwidth. On the other hand, the latter issue, the uncertainty in the traffic variation, is more complicated and so it is not easily resolved.

The deployment of MTC services over the wireless network such as LTE (Long Term Evolution) raises new challenges to manage the network resources, because the capacity of the LTE network is limited and the incumbent HTC users should not be faced with service degradation due to interference caused by addition of MTC traffic.

The first problem to the ISP (Internet Service Provider) in introducing the MTC to the LTE is to provide (dimension and/or allocate) the network resource at the AP (Access Point), because the network resource at AP is limited.

There are various ways to provide the network resource to MTC over the wireless network: from complete sharing with HTC to complete separation from HTC, where partial sharing lies in between the two extremes.

There are lots of works on the issues in the provision of resources and quality of service (QoS) to multiple types of MTC itself to the current Internet (See [1, 2, and 3, and the references therein]).

When it comes to standardization activity for the provision of QoS to MTC, 3GPP proposes an ACBS (Access Class Barring Scheme) for various types of user equipments (UEs) for MTC, which aims to enhance the QoS of MTC and to avoid overload in RAN (Radio Access Network)[4]. However, no explicit scheme is presented there.

※ This research is financially supported by Changwon National University in 2014~2015.
•° First and Corresponding Author : Dept. of Information & Communications Engineering, Changwon National University, hoony@cwnu.ac.kr, 종신회원
논문번호 : KICS2015-03-046, Received March 6, 2014; Revised May 26, 2015; Accepted June 4, 2015

Jin et al. investigated various types of network architecture and QoS issues for Internet of things (IoT)[1]. However, there is no quantitative evaluation about the argument.

Potsch et al. investigated the influence of machine-to-machine (M2M) traffic on the LTE system to the conventional voice and video services, where they showed that the strict priority scheme for packet scheduling results in satisfactory results on the performance of the conventional HTC services[2]. The limits on this work lie in the scale of the experiments where the number of the source in conventional voice and video service is only 10 and the number of the device for MTC is not greater than 1,000, which is not true in the real field.

Lien et al. proposed a two-class packet access scheme for the MTC with deterministic and stochastic QoS classes[3]. However, we could find no work that deals with the performance of the network with a massive number of MTC and HTC at the same AP in that work or elsewhere.

Note also that the standard activities do not care about the implementation of the resource allocation or dimensioning for a network that deals with HTC and MTC. Instead, they propose a network architecture and enhancement of system for MTC over the network that supports the conventional HTC[5].

To the best of author's knowledge, we could find no work that investigates the performance of IP network that accepts both HTC and MTC at the same time. Therefore, this work aims to investigate the problem of the conventional resource allocation schemes in the Internet where a massive number of both HTC and MTC coexist.

It is evident that the most appropriate method to investigate the performance of the network provisioning for both HTC and MTC at the current Internet is to simulate a network environment by utilizing the well-established simulation tools such as an OPNET or NS-2 by assuming generic characteristics of the network and the specification of the services that are associated with it.

However, it is difficult or almost impossible to simulate the experimental environment for the network system with MTC and HTC, because of the following two reasons: First, there exists no input model that generally covers the traffic profile of MTC, which is even more non-deterministic and amorphous, because it depends heavily on the devices that are connected at the AP (See [6]). Second, the number of the devices in MTC is in the order of hundreds or thousands, which has a scalability problem.

To overcome the difficulty in the simulation we propose a theoretical framework to that problem. The basic platform of the theory is divided into two parts: One is a Gaussian approximation for the arrival process in MTC as well as HTC, the other is a Max-Min analysis for the performance assessment of the network resource.

The rationale for using the above-mentioned framework is as follows: The Gaussian approximation can cover the process of input traffic in most broad spectrum. The Max-Min analysis can estimate the optimistic and pessimistic performance of the network resource, whereas the conventional mean analysis can only evaluate the average performance.

This paper is composed as follows: In Section II, we review the related works concerning the network dimensioning in Internet. In Section III, we summarize the characteristics of MD and HD. In Section IV, we present the service model and mechanisms for capacity allocation in the Internet that accommodates MD and HD. In Section V, we propose a Max-Min analysis of the service models given in Section IV, which is the main contribution of this work. In Section VI, we present the result of the numerical experiment. In Section VII, we summarize this work.

Ⅱ. Related Works on Network Dimensioning

When it comes to the network dimensioning, one could classify it into two categories: offline dimensioning (OfD) and online dimensioning (OnD)[7]. The former focuses on the long-term traffic demand, whereas the latter focuses on the short-term or dynamic demand.

It is usual that OfD is used for designing and/or reconfiguring the network capacity and OnD is used for the congestion control and/or resource allocation.

There are numerous methods to classify the OfD and OnD (See [7]). However, we argue that OfD is classified into deterministic and statistical, whereas OnD is classified into statistical and stochastic, which is summarized as follows:

First, let us consider OfD. Deterministic approach assumes the allocation of the network resource to customers via a worst case (PVD, peak value dimensioning) or mean case (MVD, mean value dimensioning). Note that PVD has to be used in a network with strict QoS requirement. Here, there is a trade-off: PVD is too much costly and MVD is too much risky in guaranteeing QoS (*In this work QoS is defined by the amount of network resource that is available to the traffic class*). The statistical approach in OfD assumes the allocation of the network resource to customers based on the long-term statistics of the customers' demand such as mean and variance.

Now let us consider OnD. Statistical approach in OnD assumes the allocation of the network resource to customers based on the short-term statistics of the customers' demand such as mean and variance. On the other hand, stochastic approach in OnD uses the queuing or time-series model to determine the capacity of the network resource. Even though the stochastic approach may reflect the time-varying nature of the dynamic traffic demand, it is NP-hard to faithfully model the multi-variate input and diverse demands of QoS from different applications that reside in the same network.

Summarizing our discussion about OfD and OnD, one can find the following fact. When it comes to accuracy in the model, we have the following order: *stochastic* > *statistical* > *deterministic*. When it comes to the cost of network resource, we have the following order: *stochastic* < *statistical* < *deterministic*.

Here, we argue that the statistical approach is a compromise between the two extremes of deterministic and stochastic approaches, and this work focuses on the statistical approach. A detailed discussion on the algorithm is given in Section IV.

Ⅲ. Human Data and Machine Data

Let us first summarize the characteristics of HD and MD. In [8] Lee argued that there are generic characteristics of MD that are different from HD: (1) MD are generated from a large number of devices. (2) Packets in MD are small-sized. (3) Upstream traffic is much greater than downstream traffic. (4) Packet transmission is infrequent and bursty. (5) Packet has flexible time-requirement. (6) Access to the network is group-based.

On the other hand, HD have the following attributes: (1) HD are generated from a moderate number of smartphone. (2) Packet size is moderate and large. (3) Downstream traffic is much greater than upstream traffic. (4) Packet transmission is frequent and non-bursty. (5) Packet has inflexible time-requirement. (6) Access to the network is flow-based.

One can see from the above discussion that MD has almost opposite characteristics as compared with HD. Among them, massive number of machine devices and burst transmission of the machine data imposes serious problem on the performance of network resource.

On the other hand, human data can be classified into interactive data (ID) and non-interactive data (NID). Voice conversation and online-game belong to ID, whereas web-related applications and streaming service belong to NID.

As to the machine data, the data type is classified in various manners, which is dependent on the application. In this work let us assume a smart grid which is considered to be the most widely penetrated service for the near future MTC.

As to the type of machine data in smart grid there exist two traffic types. One is the usual operational data (UOD) and the other is sporadic event-driven data (SED)[9].

As to the UOD, the device conditions and power quality have to be measured in a periodic manner, which incurs a sustainable provision of the network resource to UOD. On the other hand, the transfer of SED is triggered by detected failure in the power system, which requires immediate provision of the network resource to SED.

Here we have a problem: When UOD and SED compete for a network resource, what type of data has to be served first? In our opinion, SED has to preempt UOD so that any failure can be addressed immediately.

From the characteristics of HD and MD, one can find that some packets from HD can be lost or delayed without a serious problem in the experienced QoS, whereas MD (especially for smart grid) do not permit any failure in the transfer of packets, otherwise system failure occurs and this is a serious problem to the wide-scale society. That is to say, HD can tolerate some level of performance degradation, which is usual in the current Internet. On the other hand, MD is intolerant about the performance degradation. There exists a dilemma: the incumbent HD has to yield the network resource to the newly added MD.

To overcome this problem, the network operator has to devise a new service model for the MD that competes for the limited network resource with HD. However, we could find no work that deals with this issue. In the following section we consider a number of service models concerning this problem.

Ⅳ. Service Model and Resource Allocation

The most general service model in the current Internet is BE (Best Effort), where the network capacity is limited and packets from every application compete for a shared network resource regardless of the type of the service. A network that adopts BE service model is usually dimensioned based on the mean rate of the input traffic[10,11]. In [12] we have proposed a link dimensioning model for MTC with better than BE (BTBE), where the link capacity is dimensioned in a manner that allows a statistical variation of the input traffic.

However, it is envisioned at this stage that BTBE is not introduced to the Internet for the provision of MTC, because it requires a much higher cost than BE. Therefore, we conjecture that BE is the realistic service model.

Let r_{HD} and r_{MD} be the mean required service rate (unit: packets per service opportunity (SO)) of HD and MD, respectively. Let s_{HD} and s_{MD} be the average packet size (unit: byte) of HD and MD, respectively. Then, the input rate (λ_i, where i=*HD* or *MD*) of each data can be represented as follows:

$$\lambda_i = r_i \times s_i, where\ i = HD \text{ or } MD \quad (1)$$

When HD and MD compete for the same resource in a network with BE service model, the packets are served in first-in-first-out (FIFO) manner. In this case, one can't expect a differentiated level of service between HD and MD.

If the packets from HD and MD have to be treated with different levels of QoS over the BE service, the ISP has to prepare an appropriate resource allocation model for HD and MD.

When it comes to the resource allocation model, one can assume three typical types: shared service (this is a plain BE model), dedicated service (total separation of HD and MD over BE model), and hybrid of dedicated and shared service (a compromise between the extremes), each of which is described below.

4.1 Shared service

In the shared service (SS), the total service capacity of the network resource in an SO is shared by both MD and HD. This corresponds to a BE service model in the Internet. SS is usually called as a *co-channel service* in wireless network.

The network resource is a two-dimensional area that is composed of time slots and frequency band, which is typical in the orthogonal frequency division multiple access (OFDMA) that is adopted in the current broadband wireless access network such as Wi-Fi, LTE, and WiMAX.

The amount of service rate (let it be C) in an SO that is required to serve both the MD and HD traffic is given as follows:

$$C = \lambda_{MD} + \lambda_{HD} \quad (2)$$

This service model is in line with the current BE service model, so it operates well when the QoS requirements from MD and HD are the same or when the offered load to the system is low or moderate.

The main drawback of the SS is the interference between MD and HD when the volume of the traffic from each type of data varies with uncertainty. For example, when the volume of traffic from the MD (or HD) increases in an SO, the packets from HD (or MD) may not be served in that SO. This results in a severe degradation of QoS to the interfered data. Therefore, SS model is not suited when a differentiation for the packet service is needed for MD.

4.2 Dedicated service

In a dedicated service (DS), the network resource is separated into that of MD and HD, where the network resource allocated to MD is orthogonal to that of HD, and vice versa. Therefore, no interference exists between MD and HD. DS is usually called as an *orthogonal-channel service* in wireless network.

In DS, a portion of the service capacity C of the network in an SO is dedicated to MD (C_{MD}) and HD (C_{HD}) in the following manner:

$$\begin{aligned} C_{MD} &= \frac{\lambda_{MD}}{\lambda_{MD}+\lambda_{HD}} C \\ C_{HD} &= \frac{\lambda_{HD}}{\lambda_{MD}+\lambda_{HD}} C \end{aligned} \tag{3}$$

The advantage of DS is that each type of data has an exclusive guarantee of SO without interference from the other.

The main drawback of DS is low utilization when the traffic volume of one or the other data is dynamically varying, because it does not allow for MD (or HD) to utilize the resource that is not consumed by HD (or MD). Therefore, this model is suitable when the traffic volume of each type of data is very static.

SO is defined to be an opportunity for a flow to transfer the packet, and it is allocated at every frame interval.

4.3 Dedicated and shared service

It is usual that the importance of the data between MD and HD is not necessarily the same. It is envisioned that MD is considered to be tighter in the requirement of QoS than HD, because a large part of MD is generated from mission critical utilities such as smart grid, e-health, environment monitoring, etc. Therefore, the NO has to allocate the network resource to MD with high probability over HD when MD and HD compete for a network resource. This work assumes that the priority is set to MD.

The priority in the Internet is realized in various levels: priority in packet scheduling and priority in resource allocation. One can adopt the two in separate or in combination, which depends on the policy of the NO.

This work considers a combination of the two, the basic philosophy of which is given as follows: First, the network resource is allocated to MD and HD in a long term manner. Second, the access of packet from MD and HD is enforced, the detail of which is given as follows:

First, the network resource is allocated to each type of data in the long term scale. Assume that the network resource C is divided into two areas: area 1 with capacity C_D, which is dedicated to MD and area 2 with capacity C_S=C-C_D, which is shared by both MD and HD. This scheme is called a DSS (Dedicated and Shared Service).

The size of each area is determined by the arrival rate of packets from each type of data, which is given as follows:

$$\begin{aligned} C_D &= \frac{\lambda_{MD}}{\lambda_{MD}+\lambda_{HD}} \equiv \alpha C \\ C_S &= C - C_D = \frac{\lambda_{HD}}{\lambda_{MD}+\lambda_{HD}} C \equiv (1-\alpha) C \end{aligned} \tag{4}$$

Note that, even though an SO is allocated to each area based on the mean rate of each data, the real arrival of data is variable. Due to the uncertainty of the input process, the instantaneous packet arrival rate from MD or HD can exceed the capacity of C_D or C_S. Because we have assumed a service priority to MD over HD, there must be some resolution that addresses the variation of the input rate from MD, whereas no remedy has to be done for HD.

As a resolution, the second scheme is introduced, where the network resource allocated to HD is adjusted in a short term scale. The main point of the second scheme is to allow that a certain amount of the network resource in area 2 is exclusively used by MD packets.

There exist lots of algorithm to realize the above service, one of which is to allow $a\times 100\%$ of the MD packets enter area 1 and the remaining packets enter area 2. HD packets can only enter area 2. By doing that, one can realize a sustained guarantee of QoS to MD even in case of high variable traffic volume in short term scale.

Let η_k be the mean arrival rate of packets to area *k*, where *k=1* or *2*. Then, we have the following formulae for η_k (*k=1* or *2*).

$$\begin{aligned}\eta_1 &= \alpha\lambda_{MD} \\ \eta_2 &= (1-\alpha)\lambda_{MD} + \lambda_{HD}\end{aligned} \quad (5)$$

V. Max-Min Analysis

In Section IV, we have presented typical schemes for a service model about MD and HD at AP, where a BE model is assumed and the network resources is allocated in terms of the mean rate.

However, the instantaneous arrivals of packets from MD and HD behave not in deterministic but in stochastic manner, where the input rate from MD and HD fluctuates above or below the mean rate at every SO. This necessitates a model that reflects the randomness in the input rate of the packets into the resource allocation.

In [12] we have argued that the arrival of packets from a massive number of customers follow the central limit theorem. Therefore, one can assume that the arrival of packets from MD and HD follow normal distributions with mean and variance, which is given by (λ_i, σ_i^2), where *i= MD* or *HD*. The arrival of packets from MD and HD fluctuates in the range $[\lambda_i - k\sigma_i, \lambda_i + k\sigma_i]$ (*k>0* and *i= MD* or *HD*)[13].

It is known that some MTC applications have higher QoS requirements than normal data services[6]. Therefore, the traffic characteristics such as the volume and fluctuation of MD have an important role on the performance of AP.

There exist two extreme cases of fluctuation in the input rate of MD or HD: maximum and minimum fluctuation, where the input rate ranges between $\lambda_i + k\sigma_i$ (maximum fluctuation, MaxF) and $\lambda_i - k\sigma_i$ (minimum fluctuation, MinF), where *i= MD* or *HD*.

Now let us compute the amount of network resource consumed by each service model.

5.1 Shared service

For a shared service model, the packets from MD and HD compete for the SO of capacity *C*. The range of the offered load from HD and MD is given as follows:

$$\begin{aligned}\rho_{HD} &= [\lambda_{HD} - k\sigma_{HD}, \lambda_{HD} + k\sigma_{HD}] \equiv [\rho_{HD}^-, \rho_{HD}^+] \\ \rho_{MD} &= [\lambda_{MD} - k\sigma_{MD}, \lambda_{MD} + k\sigma_{MD}] \equiv [\rho_{MD}^-, \rho_{MD}^+]\end{aligned} \quad (6)$$

Note that there are four cases of extreme mixing in the offered load, which is denoted as follows:

$$\begin{aligned}\rho^{--} &= \rho_{HD}^- + \rho_{MD}^- \\ \rho^{\mp} &= \rho_{HD}^- + \rho_{MD}^+ \\ \rho^{\pm} &= \rho_{HD}^+ + \rho_{MD}^- \\ \rho^{++} &= \rho_{HD}^+ + \rho_{MD}^+\end{aligned} \quad (7)$$

Now let us investigate the relationship between the offered load and the provided service for each case. First, let us consider ρ^{--}, which is given by

$$\rho^{--} = \lambda_{HD} + \lambda_{MD} - k(\sigma_{HD} + \sigma_{MD}) \quad (8)$$

Note that $\rho^{--} < C$, so there is no deficiency in the network resource. Second, let us consider $\rho^{\mp}$, which corresponds to

$$\rho^{\mp} = \lambda_{HD} + \lambda_{MD} + k(\sigma_{MD} - \sigma_{HD}) \quad (9)$$

In this case, there is no deficiency in the network resource for serving the data from both HD and MD when $\sigma_{MD} \leq \sigma_{HD}$. Instead, there is a surplus in the network resource that is not consumed by both HD and MD, which is denoted by $\epsilon_{Tot}^{\mp} = k(\sigma_{HD} - \sigma_{MD})$. When $\sigma_{MD} > \sigma_{HD}$, the deficiency in MD is given as $\delta_{MD}^{\mp} = k(\sigma_{MD} - \sigma_{HD})$. Third, let us consider $\rho^{\pm}$, which corresponds to

$$\rho^{\pm} = \lambda_{HD} + \lambda_{MD} + k(\sigma_{HD} - \sigma_{MD}) \tag{10}$$

Note that there is no deficiency in the network resource for serving the data from both HD and MD when $\sigma_{MD} \geq \sigma_{HD}$. Instead, there is a surplus in the network resource that is not consumed by both HD and MD, which is denoted by $\epsilon_{Tot}^{\pm} = k(\sigma_{MD} - \sigma_{HD})$. When $\sigma_{MD} < \sigma_{HD}$, the deficiency in HD is given as $\delta_{MD}^{\pm} = k(\sigma_{HD} - \sigma_{MD})$ Finally, let us consider ρ^{++}, which corresponds to

$$\rho^{++} = \lambda_{HD} + \lambda_{MD} + k(\sigma_{HD} + \sigma_{MD}) \tag{11}$$

The deficiency in HD and MD is inevitable in this case because $\rho^{++} > C$, which is equal to $\delta_{Tot}^{++} = k(\sigma_{MD} + \sigma_{HD})$. There is no surplus in the network resource.

Summarizing the above discussion, we arrive at the following observation: First, the deficiency in the network resource is in the range of $[0, k(\sigma_{HD} + \sigma_{MD})]$. Second, the deficiency in the network resource depends on the variance of the offered traffic, which is observed from the second and third cases in the above discussion.

In practice, it is usual that $\sigma_{MD} \leq \sigma_{HD}$ (See [6]), which implies that MD do not have to worry about the deficiency in the network resource when the offered traffic from HD is light. However, the deficiency in the network resource to MD is not negligible when the offered traffic from HD is heavy.

5.2 Dedicated service

For DS, the packets from MD and HD can consume C_{MD} and C_{HD}, respectively, at each SO. Therefore, it is straightforward to estimate the surplus or deficiency in the network resource. Note that the packets from MD and HD require $[\rho_{MD}^{-}, \rho_{MD}^{+}]$ and $[\rho_{HD}^{-}, \rho_{HD}^{+}]$, respectively, in each SO. However, the provisioned resource is λ_{MD} and λ_{HD} at each dedicated area, respectively. In this case, the surplus and deficiency in the network resource are orthogonal between MD and HD, which equals to $\Phi_{MD} = k\sigma_{MD}$ and $\Phi_{HD} = k\sigma_{HD}$ for MD and HD, respectively.

5.3 Dedicated and shared service

For DSS, the priority in the packet service lies in MD, so that the Max-Min analysis is given for the two extreme cases of heavy MD $(\lambda_{MD} + k\sigma_{MD})$ and light MD $(\lambda_{MD} - k\sigma_{MD})$ traffic, which is given as follows: For the heavy MD traffic, the packets from MD consume $\alpha(\lambda_{MD} + k\sigma_{MD})$ and $(1-\alpha)(\lambda_{MD} + k\sigma_{MD})$ of capacity from area 1 and area 2, respectively, in each SO. In this case, the capacity of network resource in area 2 that can be consumed by HD in each SO is reduced to $(1-\alpha)(C - k\sigma_{MD})$.

First, let us compute the surplus in the network resource that is not consumed by HD.

Since the offered traffic of HD is in the range [$\lambda_{HD} - k\sigma_{HD}, \lambda_{HD} + k\sigma_{HD}$] there is a surplus when $(1-\alpha)(C - k\alpha_{MD}) - \lambda_{HD} + k\sigma_{HD} > C$

Otherwise, there is no surplus.

The Max-Min range of deficiency in HD is given as follows:

$$\varphi_{HD}^{heavy\_MD} = [\alpha\lambda_{HD} - k\sigma_{HD} - (1-\alpha)(\lambda_{MD} - k\sigma_{MD}), \alpha\lambda_{HD} + k\sigma_{HD} - (1-\alpha)(\lambda_{MD} - k\sigma_{MD})]$$

For the light MD traffic, the packets from MD consume the network resource in an amount of $\alpha(\lambda_{MD} - k\sigma_{MD})$ and $(1-\alpha)(\lambda_{MD} - k\sigma_{MD})$ from area 1 and area 2, respectively, in each SO. The amount of network resource that can be used by HD is $(1-\alpha)(C + k\alpha_{MD})$.

First, let us compute the surplus in the network resource that is not consumed by HD.

Since the offered traffic of HD is in the range $[\lambda_{HD} - k\sigma_{HD}, \lambda_{HD} + k\sigma_{HD}]$, there is a surplus when $(1-\alpha)(C + k\sigma_{MD}) - \lambda_{HD} + k\sigma_{HD} > C$.
Otherwise, there is no surplus.

The Max-Min range of deficiency in HD is given as follows:

$$\varphi_{HD}^{light\_MD} = [\alpha\lambda_{HD} - k\sigma_{HD} - (1-\alpha)(\lambda_{MD} + k\sigma_{MD}), \alpha\lambda_{HD} + k\sigma_{HD} - (1-\alpha)(\lambda_{MD} + k\sigma_{MD})]$$

Note that MD does not experience deficiency in any case.

Up to now we have investigated the Max-Min range of the network usage for the various mixings of HD and MD traffic, and we have shown the surplus and deficiency of the network resource for each case.

However, it is evident that the worst-case situation in the traffic volume plays an important role in the evaluation of the network performance. So, let us focus the case of heavy traffic from both MD and HD.

The deficiency of network resource seen from HD and MD differs from the service model, which is summarized in Table1.

From the above results, one can find the following two facts: First, the amount of the network resource that is deficient in HD and MD depends on the nature of the associated data. Second, the expected QoS of the HD and MD is in the order of DSS > DS > SS, which means that DSS has the most superior performance.

Table1. Resource deficiency

Service model	Resource deficiency	
	HD	MD
SS	$k(\sigma_{HD} + \sigma_{MD})$	$k(\sigma_{HD} + \sigma_{MD})$
DS	$k\sigma_{HD}$	$k\sigma_{MD}$
DSS	$(1-\alpha)k\sigma_{MD}$	0

Ⅵ. Numerical Result and Discussion

Let us investigate the performance of the proposed analytical model, which is based on the resource deficiency that has been shown in Table1. As one can find from Table1, the performance of the AP depends on the following parameters: For SS and DS, k and $\sigma_i (i = MD \text{ or } HD)$, both of which indicate the fluctuation of the traffic from both MD and HD. For DSS, α is added to them, which is a control parameter for the resource allocation between MD and HD. Let us assume those parameters as shown in Table 2.

In Table 2 we assume $\overline{\omega}$ (relative STD, the relative standard deviation of HD as compared with that of MD) as a variable, because $\overline{\omega}$ is considered to be the crucial factor that determines the performance of the AP. For simplicity, σ_{MD} is normalized to 1. The boundary of the assurance for the bandwidth is assumed to be 95.45% of the total variation. We can assume three cases for the proportion of MD traffic depending on the role of an AP: low MD and high HD ($\alpha = 0.1$), medium MD and medium HD ($\alpha = 0.5$), and high MD and low HD ($\alpha = 0.9$).

Fig.1 shows the resource deficiency of MD as a function of $\overline{\omega}$. In Fig.1, dss_MD, dds_MD, and ddss_MD indicates the deficiency of MD for SS,

Table 2. System parameters

parameter	value
$\overline{\omega} = \sigma_{HD}/\sigma_{MD}$	Input variable
k	2 (95.45% assurance)
α	0.1, 0.5, or 0.9

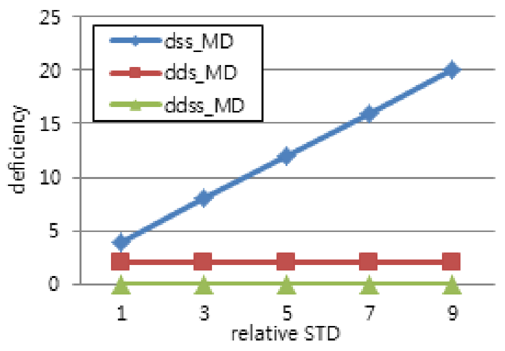

Fig. 1. Resource deficiency of MD

DS, and DSS, respectively. The result is independent from the value of α (See Table1), so we do not indicate the value of α in the figure.

From Fig.1, we found the following facts: The resource deficiency for MD increases as the relative STD increases when the packet service model is SS. When DS is used, the resource deficiency depends only on the traffic characteristics of MD itself. No deficiency of resource in MD occurs for DSS. Therefore, the performance of the service model is superior in the order of DSS > DS > SS.

Fig.2 shows the resource deficiency of HD for the same parameters, where dss_HD, dds_HD, and ddss_HD indicates the deficiency of HD for SS, DS, and DSS, respectively. The number inside the parenthesis in ddss_HD indicates the value of α, because the deficiency depends on the value of α.

From Fig. 2, we found the following facts: First of all, the resource deficiency for HD increases as the relative STD of the traffic volume increases when the packet service models are SS and DS, where SS is more sensitive than DS. When it comes to DSS, the deficiency of resource in HD is constant, where its value is very small compared with those of SS and DS.

Note also from Fig.2 that the deficiency of resource in HD is not so sensitive to the proportion of MD among the total traffic when the resource is allocated in DSS, which has a very important implication in the dimensioning of the link. That is to say, there is no serious problem of resource deficiency for HD if DSS is introduced in an AP when MD is added to HD.

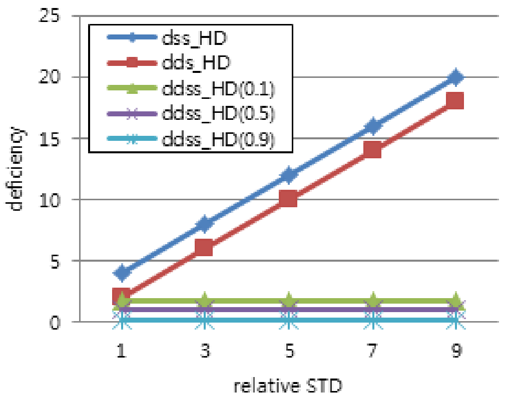

Fig. 2. Resource deficiency of HD

Summarizing our discussion about the performance of the resource allocation in AP that accommodates the MD and HD, the performance of the service model is superior in the order of DSS > DS > SS, which follows the same pattern as that of MD.

In conclusion, we recommend DSS as the most appropriate resource allocation scheme for MD and HD in AP.

Ⅶ. Conclusion

In this work we have proposed a framework for the performance assessment of an AP in LTE network that accommodates both MD and HD. To that purpose, we investigated the characteristics of the traffic from MD and HD.

Typical service models and resource allocation methods in LTE for MD and HD have been investigated.

After that we drew out a boundary for the surplus and deficiency of network resource about MD and HD by using a Max-Min analysis.

Finally, we presented the validity of the proposed theoretical framework via a numerical experiment, from which we have found that the performance of the DSS scheme is superior to that of the SS and DS schemes.

In conclusion, we argue that a systematic analysis on the performance of the AP has to be done in order to provide MTC service to the current LTE network by investigating the traffic characteristics of MD and the mixing of the traffic between MD and HD, via which the QoS of MD as well as HD can be provided to a desired level.

Our future work includes the consolidation and sophistication of the proposed framework by applying real-field data about MD and HD at AP of the LTE network.

References

[1] J. Jin, J. Gubbi, L. Tie, and M. Palaniswami, "Network architecture and QoS issues in the internet of things for a smart city," in *Proc. ISCIT*, pp. 956-961, Gold Coast, Oct. 2012.

[2] T. Potsch, S. N. K. Marwat, Y. Zaki, and C. Gorg, "Influence of future M2M communication on the LTE system," *IFIP WMNC*, pp. 1-4, Dubai, Apr. 2013.

[3] S-Y Lien, et al., "Massive access management for QoS guarantees in 3GPP M2M communications," *IEEE Commun. Lett.*, vol. 15, no. 3, pp. 311-313, Mar. 2011.

[4] 3GPP TR 37.868 v11.0.0 (2011-09); Technical Specification Group Radio Access Network; *Study on RAN Improvements for Machine-type Communications*; (Release 11).

[5] S. Hussain, et al., "Recent trends in standards related to the Internet of things and machine-to-machine communications," *J. Inf. Commun. Converg. Eng.*, vol. 12, no. 4, pp. 228-236, Dec. 2014.

[6] D. Boswarthick, et al., *M2M communications: a systems approach*, Wiley, 2012.

[7] S. Yang and F. A. Kuipers, "Traffic uncertainty models in network planning," *IEEE Commun. Mag.*, pp. 172-177, Feb. 2014.

[8] H. Lee, "The seven characteristics of M2M communications: Observations and implications," *J. Changwon National Univ.*, vol. 2, Dec. 2013.

[9] W. Sun, et al., "Quality of service networking for smart grid distribution monitoring," *IEEE Int. Conf. Smart Grid Commun.*, pp. 373-378, Gaithersburg, 2010.

[10] R. S. Cahn, *Wide area network design*, Morgan Kaufmann Publishers, Inc., 1998.

[11] H. Lee, *Internet Design, Lecture note*, Changwon National Univ., 2015.

[12] H. Lee, "Link dimensioning for multiplexed M2M traffic," *J. KICS*, vol. 39, no. 1, Jan. 2014.

[13] H. Lee, *Lectures on Probability 2.0*, Hongreung Publishing Co., 2014.

이 훈 (Hoon Lee)

1984년 : B.E. from Kyungpook Nat'l Univ., Korea

1986년 : M.E. from Kyungpook Nat'l Univ., Korea

1996년 : Ph.D. from Tohoku Univ., Japan

1986년~2000년 : KT R&D Center

2001년~Present : Changwon National Univ.

2005년~2006년 : Univ. of Missouri-Kansas City, USA

2011년~2012년 : Marquette Univ., USA

<Research fields> Design of Internet