On Modeling Ubiquitous Cloud: Estimation of Traffic

Shuji Kawasaki, Masakazu Higuchi and Hitomi Murakami
Seikei University
Department of Information and Computer Science
Kichijoji-Kita, 121-35 Musashino
Japan
{shuji1113, m-higuchi, hi-murakami}@gmail.com

Abstract: The Ubiquitous Cloud is a concept of large-scale information service network as a social infra-structure. It is featured by real-world context information extraction, user information profiling and self-cofiguration/-control of network. The objective of the research is to evaluate the network traffic theoretically and thus give a guideline of network design. In this paper, we especially consider those traffic factors that are related to movement of mobile users and rea-time services, and thus discuss some necessary requirements for the network specification.

Key–Words: context-awareness, profiling, self-reconfiguration, spatio-temporal dynamics, long-range dependence, self-similarity, elephant-mice flow.

1 Introduction

The Ubiquitous Cloud (UC) is the framework of a ubiquitous network concept that has been advocated since 2003 aiming at providing a theoretical and practical basis of a prospective social infra-structure. The R&D of the concept is conducted in the Ubiquitous network control and administration (Ubila) project of Japan. The project members are experts in the field from companies and universities.

UC is a information service network, as in Fig. 1, with the following features:

- Autonomic extraction of various context information from the real world and coordination of an appropriate service for a user
- Keeping to provide a better service by user information profiling
- Self-reconfiguration of the network

The supposed contents of the context information include the matter of food, clothing and shelter, crime/disaster prevention and rescue, meidical/welfare /nursing services, vehicles and ITS (intelligent transport system), economics and business, amenity and favor, etc. The context information of a user in the real-world is detected at any time by sensors deployed everywhere and sent to a server called cloud, which is the brain and controller of the network. The cloud then forms an appropriate service information for the user according to the context or application and sends it to a nearby actuator to provide a real-world service. We call a pair of sensor and actuator a *node*.

Technical elements close to user services level suppose context-awareness /-modeling and location-awareness. These are new themes in the relevant field. Especially for the first one, R&D has just begun and useful results has not obtained yet. In a physical or technological level, on the other hand, suppose IPv6, Heteroscedastic multiplexing, Ad Hoc network, real-time scheduling, etc.

For such large-scale infrastructure network UC, it is important to assess its network traffic theoretically and give a designing guideline. The objective of the research is obtain the theoretical assessment. Especially we want to know

- how much network capacity is necessary
- how much spatial density of the nodes is necessary
- how long the response time from sensing to actuation is
- how quality of information is evaluated

Here an actuation implies not only the service with physical operation but giving users requested information.

Among the context fields mentioned above, it may be considered that disaster rescue, medical service, care for handicapped persons, etc. are of pubic importance and they often need real-time response. In recent years, on the other hand, a lot of applications in these fields are developed so as to utilize mobile terminals as PDA. Thus a consideration of mobile users

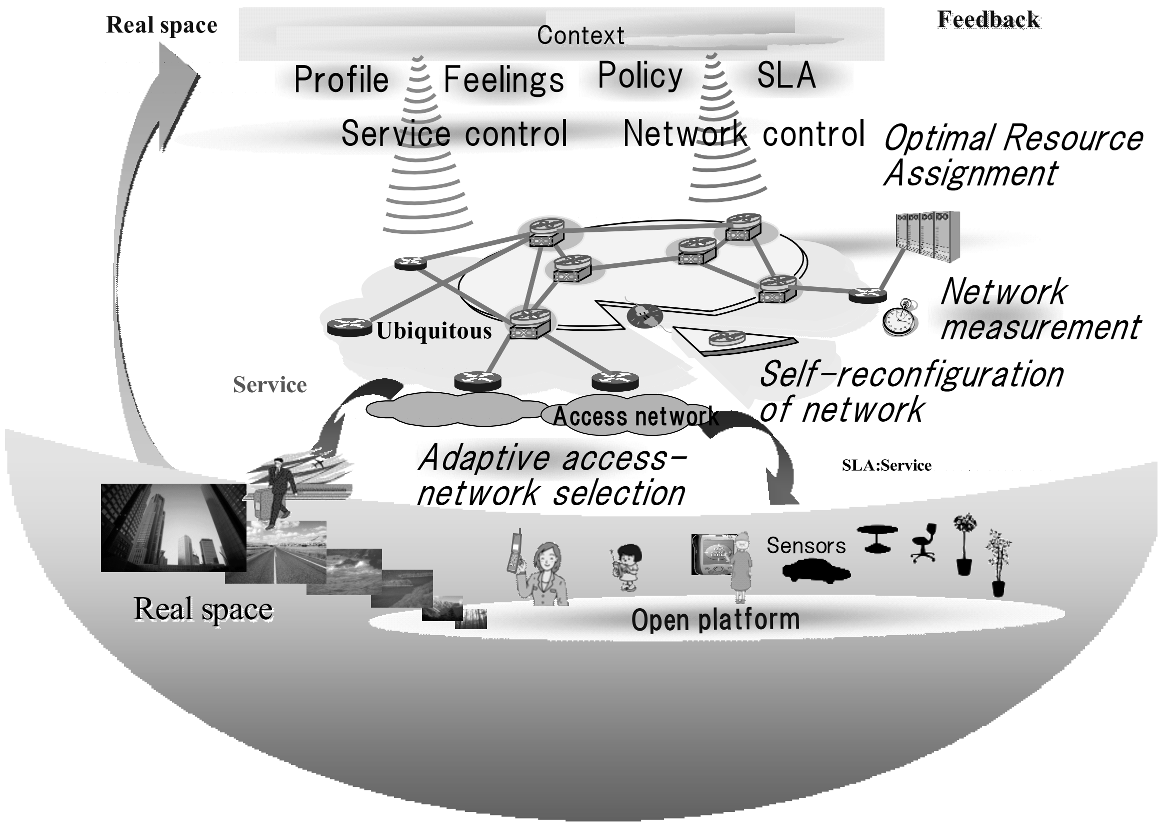

Figure 1: Ubiquitous Cloud Image

in UC may be a problem of high priority. Therefore, in this paper, we focus our attention to mobile applications that requires real-time response to clarify problems raised in the assessment.

The organization of the paper is as follows. In section 2, we survey known results so far and pick up concepts that we will follow. Section 3 is the preliminary consideration for UC traffic modeling, containing some remarks in user mobility formulation. Section 4 then considers the fundamental features that the UC should possess to seek for how to formulate the traffic and what the distinction of UC from classical queueing theory.

2 Recent Progress in Network Traffic Study

Since 1990s, a large number of research on the network traffic has been done. Their points of view are characteristics specific to protocols (TCP/IP, HTTP, FTP, etc.), network form (WAN, LAN, etc.) and application (WWW, P2P, etc.). It is pointed out commonly that the traffic has the following characteristics:

- failure of arrival process models by simple Poisson processes
- failure of traffic models by Markov processes
- self-similarity/long-range dependence; burstiness
- elephant flow

Moreover, in aggregation of a lot of user traffics, several authors reports that the aggregated traffic does not present Gaussianity, according to measurements of real traffic. Conventionally it is considered that the Gaussianity holds based on the assumption that a central limit theorem holds. However, the aggregated traffic being nonGaussian may imply that we should consider rather a noncentral limit theorem, which is sometimes the case for long-range dependent processes.

As for the burstiness, Lowen and Teich discusses the generation mechanism [12]. They give two such models called Bartlett-Lewis process and Neman-Scott process. In this study, we will consider our model regarding these characteristics as well.

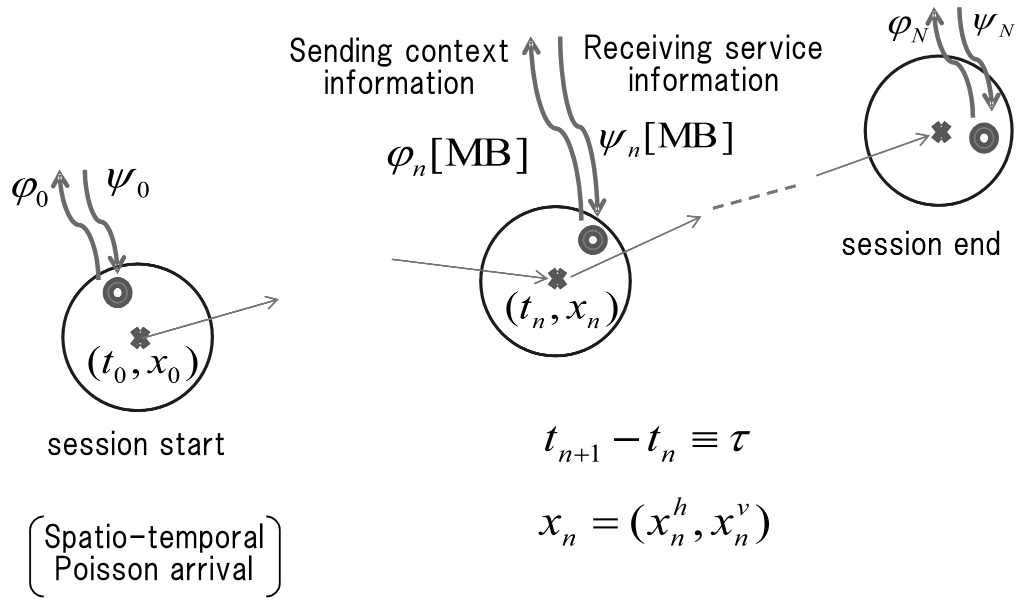

Figure 2: Basic Movement Model of a Mobile User

The original point of our model is that it is aiming at modeling of spatio-temporal dynamics of the network. In the conventional traffic model based on queueing theory, user traffics accessing to finite servers are aggregated; here the interest is in the system behaviors with respect to time as stochastic processes. Unfortunately the behaviors with respect to spatial variables are not involved. Our research considers the spatial behaviors as well as time from the following point of view.

- sessions of mobile users
- applications that involves location-awareness–disaster information, ITS, events, etc.
- consideration of spatial hop processes in Ad Hoc connection

3 Preliminary Remarks: Modeling Mobile Users

Let us consider the movement model as in Fig. 2. It presents that a mobile user starts a session at $t = t_0$ accessing a nearby node and finishes at a random time $t = t_N$. The user sends a context information at $t = t_n$ with rate φ_n[MB/s] and receives in response a service information real-time with rate ψ_n[MB/s] for $n = 1, \cdots, N$. It may be convenient to consider $\Phi_n = \max(\varphi_n, \psi_n)$ since

- sometimes the capacity of send and receive nodes are made the same
- even though the capacities are different, one can estimate a necessary capacity by a larger one.

For the sake of simplicity, we take a uniform non-random sampling rate along time, so that $t_{n+1} - t_n = \tau$ [s]. This τ may be either greater or smaller than 1, according to applications. For example, in a real-time critical system like a vehicle, that involves a faster movement than a man, τ should be smaller than 1. In a system that considers human movements, on the other hand, τ can be larger than 1.

For the first connection in a session, we assume that the connection itself obeys Poisson distribution. It is the aggregated traffic that does not obey the Poisson distribution, and we should note that the connection itself can be modeled as Poisson arrivals like BL process or NS process [12].

Fig. 2 presents that at each $t = t_n$, user position is located in a corresponding disc which is the covered-area by the node at the center of the circle. We sometimes identify the position of user and the node for the sake of easy.

It may be considered that there are a lot o mobile users like this. Let us denote the whole network region $D \subset \mathbb{R}^2$ and let D be covered by subregions D_l: $D = \cup_l D_l$. Here both setting of D_l being disjoint or joint may be possible. In this case, let us assume that in each D_l,

- new sessions start with the ratio λ_l[person/$s \cdot m^2$]
- existing session terminate with the ratio μ_l[person/$s \cdot m^2$]

(see Fig 3). If one wants to follow the stability argument in queueing theory, the one may take $\lambda_l < \mu_l$. This flow-in/-out ratio for each subregion D_l reflects the local characteristics (Fig4).

In considering the sptio-temporal model as above, it may be useful to formulate the user movements based on stochastic arguments like

- distribution of user velocities
- distribution of inter-user distances
- distribution of amount of data the users send/receive at each instance or position.

Here it should be noted, however, that if one increases the random variables in the argument easily then the user movement model could be too complicated to tract theoretically. We do not hope that but a framework that does is based on theoretical analysis. In order for that it is necessary to construct the movement model in a clever way. It is recognized that the description of user's spatial attribution is an important problem [16].

Now, in order to study the relationship between the spatial density of nodes and user velocities, we may consider

- taking a time series X_n representing the position at $t = t_n$ and evaluate their N-dimensional distribution, considering it is a Markov process
- taking a 2-dimensional time series $M_n = (V_n, \Theta_n)$, where V_n is the velocity along the line and Θ_n the angle of the moving direction and horizontal or vertical axis; assuming M_n is a Markov process, we may evaluate its N-dimensional distribution.

Also, each node has its own cover area.

In case that a session lasts over several nodes, we will think of its flow-out of a node's cover area $E_{l,n} \subset D_l$ at (t_n, x_n) and flow-in to another node's cover area $E_{l',n'} \subset D_{l'}$ at (t_{n+1}, x_{n+1}). This l' can be the same as l. In this respect, we may consider the following two ways:

- A flow-in or flow-out are involved in the starts and ends of sessions in $E_{l,n}$. The session-start ratio in $E_{l,n}$ is then taken as $\lambda_l|E_{l,n}|/|D_l|$ and session-end ratio as $\mu_l|E_{l,n}|/|D_l|$, respectively.
- The flow-in ratio and flow-out ratio are taken separately with the session-start within $E_{l,n}$ and session-end within $E_{l,n}$. In this case, the session-start ratio within $E_{l,n}$ and flow-in ratio may be denoted as $\lambda_l|E_{l,n}|/|D_l|$ and $\bar{\lambda}_l|E_{l,n}|/|D_l|$ respectively, and similarly, the session-end ratio within $E_{l,n}$ and flow-out ratio may be denoted as $\mu_l|E_{l,n}|/|D_l|$ and $\bar{\mu}_l|E_{l,n}|/|D_l|$, respectively.

In the first way, the model description is simple apparently since the difficulty of modeling the flow-in/-out is hidden away. But the model is the same as the one in which there are no movements between $\{E_{n,l}\}$. We thus employ the latter.

3.1 Congestion in a single user session

Let us denote the node that a user u_0 accesses at t by $s(t)$. If the u_0 send/receive φ[MB] and if $s(t_n) \neq s(t_{n+1})$, then the amount of data sent/received at $s(t_n)$ and $s(t_{n+1})$ are both φ/τ[MB/s], as long as the session is alive, i.e. $\varphi/\tau \geq \rho$ for the node's capacity ρ[MB/s].

Next, if $s(t_n) = s(t_{n+1})$, the amount of data sent/received during the two time slots, i.e. $[t_n, t_{n+2})$ is $\min(2\tau\rho, \varphi)$ [MB]. This can be generalized as follows. We assume that the positions of nodes and users are the same, for the sake of simplicity. Let the radius of cover area of a node be ζ[m]. If the u_0 get out of the area at $n = N^*$th step and get into another area, then the N^* is given by

$$N^* = \min\left\{ N \,\middle|\, \left(\tau \sum_{n=0}^{N} v_n \cos\theta_n\right)^2 + \left(\tau \sum_{n=0}^{N} v_n \sin\theta_n\right)^2 \geq \zeta^2 \right\}. \quad (1)$$

In this case, for k_1, k_2 with $0 \leq k_1 < k_2 \leq N^*$, a requirement may be such that the maximum of amount of data does not exceed the node's capacity in some sense. A primitive form of it may be

$$\max_{0 \leq k_1 < k_2 \leq N^*} \frac{1}{(k_2 - k_1)\tau} \sum_{i=k_1}^{k_2} \varphi_i \leq C; \quad (2)$$

Or, a probabilistic form may be as

$$P\left(\max_{0 \leq k_1 < k_2 \leq N^*} \frac{1}{(k_2 - k_1)\tau} \sum_{i=k_1}^{k_2} \Phi_i \leq C \right) \geq 1 - \delta \quad (3)$$

for certain prescribed small δ, $0 < \delta < 1$. Here N^* and Φ_i are random variables.

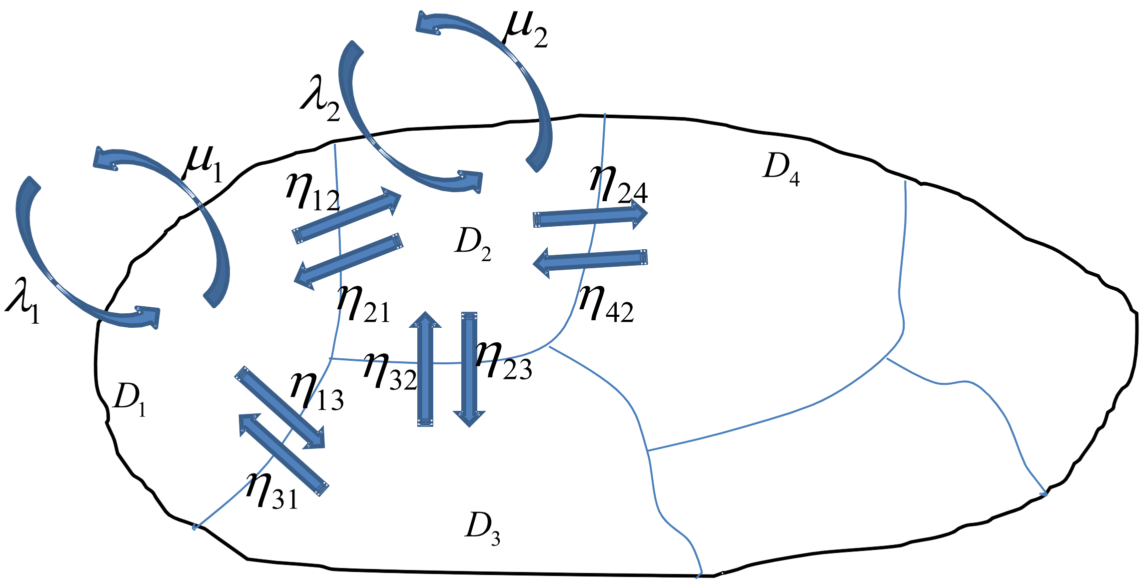

Figure 3: Area division according to coverage of cellular base stations or statistical characteristics

3.2 Congestion in Several Users

A node may be used by several users simultaneously. For this, it is immediately understood that the following requirements should be satisfied. Let us suppose that a user u_0 sends a context information of size φ[MB] from a node at $t = t_0$ and let another user $u = u_1$ sends the same data at $t = t_1$ from the same node, moving the same direction as u_0. Let the distance of the users be r[m] and verocities are the same with v[m/s]. Then

$$t_1 = t_0 + \frac{r}{v}. \tag{4}$$

If the capacity of the node is ρ[MB/s], then the time necessary to send or receive the data is φ/ρ[s]. Hence, if

$$\frac{r}{v} \geq \frac{\varphi}{\rho}, \tag{5}$$

the congestion at the node does not occur. From this itself, one might take the capacity satisfying $\rho \geq v\varphi/r$. Actually v, r, φ are, however, random variables and we may consider requirements such as

$$P\left(\frac{R}{V} \geq \frac{\Phi}{\rho}\right) \geq 1 - \delta, \tag{6}$$

where R, V and Φ are random variables corresponding to r, v and φ respectively. For some applications or positions these random variables may be set constants. For example, instead of Φ itself, the upper bound Φ^* may be used, so that random variables in (6) are only R and V.

In a more general case where users $u_1, \cdots, u_\nu$ appear in an $E_{l,n}$ simultaneously, it is sufficient for the connection not to overflow is that

$$\min_{1\leq i\leq \nu} \frac{r_i}{v_i} \geq \frac{\varphi_0 + \cdots + \varphi_\nu}{\rho}, \tag{7}$$

i.e. $\rho \geq \sum_{i=0}^{\nu} \varphi_i / \left[\min_{1\leq i\leq \nu} r_i/v_i\right]$; In a slightly more general argument the node may have a buffer, which we assume the capacity b [MB]. Then the requirement, stated in the probabilistic form, is expressed as

$$P\left(\min_{1\leq i\leq \nu} \frac{R_i}{V_i} \geq \frac{\left[\sum_{i=1}^{\nu} \Phi_i\right] - b}{\rho}\right) \geq 1 - \delta. \tag{8}$$

If we consider in each $E_{l,n}$ only those sessions that start within it and assumes the amount of data sent by each user is the same, then we can apply directly an argument of queueing theory to evaluate the loss probability at the node. The argument, called $M/M/c/c$ system, assumes Poisson arrival for each connection, exponential service time, c service stations and no waiting rooms. It is a lossy system since a customer arriving when c customers are already being served in each station then the new customer cannot enter the system.

In our model, we can think of the number c of service stations as capacity of the node. Thus the capacity of the node is c[MB/s] and each user send data to the node with the same rate s[MB/s], so that $\xi = c/s$ users can connect to the node simultaneously. Then,

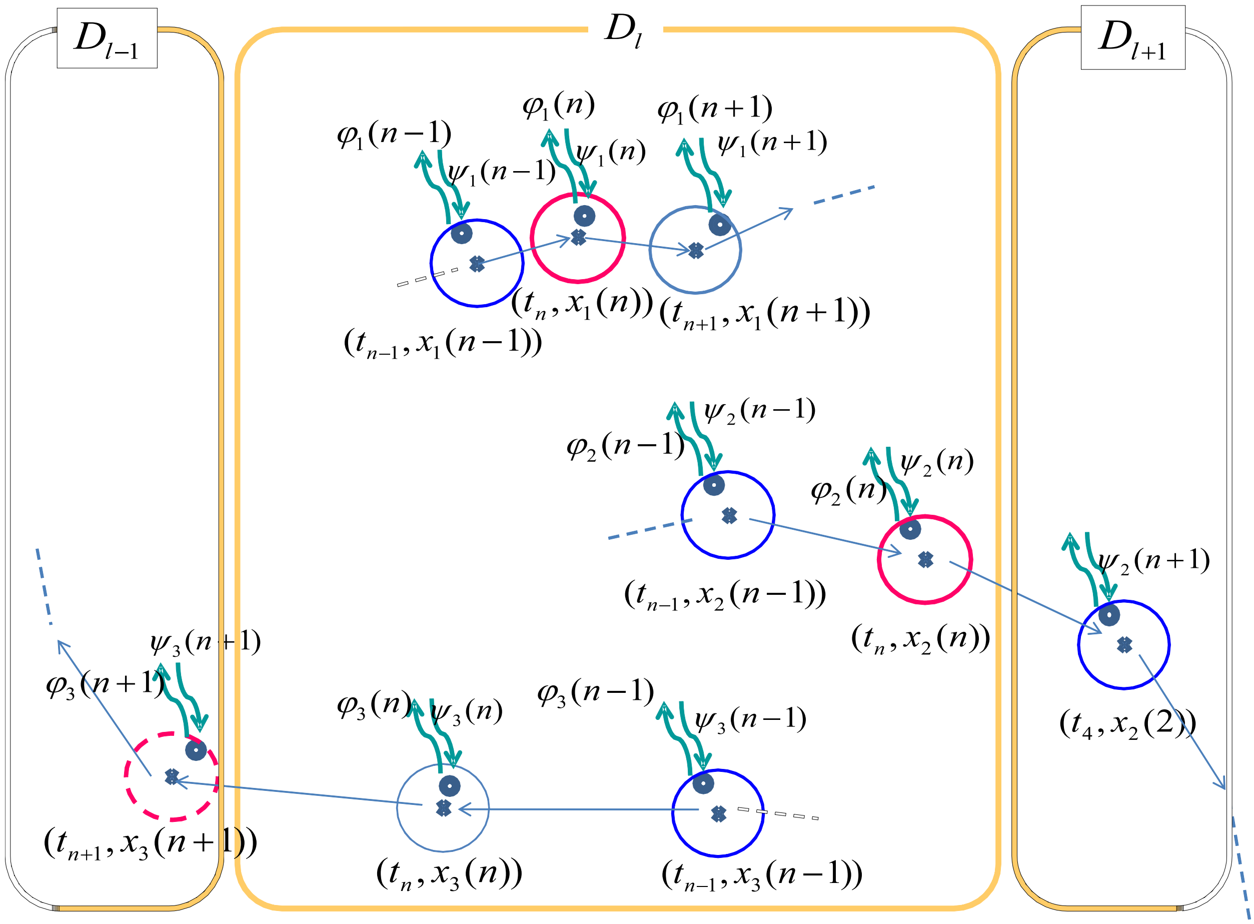

Figure 4: Flow-in/-out of sessions

the probability P_ν that ν users are connecting sessions in $E_{l,n}$ is given by [8, Section 3.7]

$$P_\nu = \frac{\rho^\nu}{\nu!} P_0, \qquad \nu \in \mathbb{N}_0, \tag{9}$$

with

$$\rho = \frac{\lambda_l |E_{l,n}|/|D_l|}{\mu_l |E_{l,n}|/|D_l|} = \frac{\lambda_l}{\mu_l} \tag{10}$$

and

$$P_0 = \left[\sum_{\nu=0}^{\xi} \frac{\rho^\nu}{\nu!} \right]^{-1}, \tag{11}$$

respectively. Also, the loss probability that a new user trying to connect when there are ξ users already is given by

$$P_\xi = \frac{\rho^\xi/\xi!}{\sum_{\nu=0}^{\xi} \rho^\nu/\nu!}. \tag{12}$$

These expressions are called Erlang's B formula.

Here it is important to recognize that we are implicitly assuming the number of users try to start a connection newly within a time slot of length τ is one. For the sake of a more realistic model, we have to take into account that several users may start connection within a time slot.

4 Towards the Modeling and Estimation of UC

In this section, we consider special features of the UC traffic model distinct from classical queuing theory. Such features may include

- spatio-temporal birth-death processes or its variants rather than temporal birth-death processes
- multiple occurrence [1][2], corresponding to multi-user connection, in a temporal point process rather than single occurrence
- the traffic process is for the amount of user traffic, and not for the number of connecting users.
- induction of clusters [1][2] by each user in each point of the temporal point processes. Especially, the clusters are such that, aggregated over the users and points, they approximate to self-similar

and LRD traffic process. In this sense, the distribution over user traffic sizes, that represents well the traffic components from mice to elephants, may be involved in the traffic modeling.

- description of user mobility by a spatio-temporal Markov process. By the spatial Markovian property, we may think of the Gibbs process [][], which is such that given the position of a user in certain boundary region of D_ls, his movement afterward is independent of the past.

Since our motivation has become clear as above, it is helpful to see a known result that may be a basis to our problem, which we do in the following subsection.

4.1 A Known Result in Jackson's Networks

In this subsection, we survey briefly the known result, called Jackson's theorem, and indicate differences from our problem. Then the Jackson's theorem will be an appropriate starting point for us and we have to seek for an extension beyond it.

The setting of Jackson's theorem is as follows [8]. We consider a network of nodes, each of which is a service facility and each with storage room for queues. Customers enter the system at various points, queue for service and upon departure from one of the nodes proceed to one of other nodes to receive there additional service.

The network consists of N nodes where the i-th node consists of m_i exponential servers each with parameter μ_i. Further, the i-th node receives arrivals from outside the system in the form of a Poisson process at rate γ_i. If $N = 1$, then it is just an $M/M/m$ system. Upon leaving the i-th node a customer proceeds to the j-th node with probability $r_{i,j}$. After completing service in the i-th node the probability that the customer departs from the network is given by $1 - \sum_{j=1}^{N} r_{i,j}$.

In order to indicate the arrival from outside and the departure to outside of the network, we designate states 0 and $N + 1$. Thus, $r_{0,i}$ is the probability that next externally generated arraival will enter the network at node i, while $r_{i,N+1}$ is the probability that a customer leaving i-th node departs from the network. $r_{0,N+1}$ is the probability that the next arrival require no service and leave immediately upon arrival. In addition, let the exponential serveice rate at node i be μ_{k_i} when the there are k_i customers at the node.

Below, we will calculate the total average arrival rate of customers to a given node. Let λ_i be the tptal average arrival rate to i-th node. Then,

$$\lambda_i = \gamma_i + \sum_{j=1}^{N} \lambda_i \, r_{j,i}, \tag{13}$$

for $i = 1, \cdots, N$. Let $\boldsymbol{k} = (k_1, \cdots, k_N)$ be number of customers in each of the nodes. Thus $\boldsymbol{k}$ is the state of the network. Then, $S(\boldsymbol{k}) = \sum_{i=1}^{N} k_i$ denotes the total number of customers in the network. Let $P_k(t)$ be the time-dependent state probabilities,

$$P_k(t) = P\Big(\text{state vector at time } t \text{ is } \boldsymbol{k}\Big). \tag{14}$$

As in the usual formulation method of queues, we can write the differential equation governing the the state probability as

$$\begin{aligned} \frac{dP_k(t)}{dt} &= -\Big[\gamma(S(\boldsymbol{k})) + \sum_{i=1}^{N} \mu_{k_i}(1 - r_{ii})\Big] P_k(t) \\ &+ \sum_{i=1}^{N} \gamma(S(\boldsymbol{k}) - 1)\, r_{0,i}\, P_{\boldsymbol{k}(i-)}(t) \\ &+ \sum_{i=1}^{N} \mu_{k_i+1}\, r_{i,N+1}\, P_{\boldsymbol{k}(i+)}(t) \\ &+ \sum_{\substack{i,j=1 \\ i \neq j}}^{N} \mu_{k_i+1}\, r_{j,i}\, P_{\boldsymbol{k}(i,j)}(t), \end{aligned} \tag{15}$$

where $\boldsymbol{k}(i-)$ equals $\boldsymbol{k}$ except for its i-th component with the i-th component is $k_i - 1$ and $\boldsymbol{k}(i+)$ equals $\boldsymbol{k}$ except for its i-th component with the i-th component is $k_i + 1$. Also, $\boldsymbol{k}(i, j) = \boldsymbol{k}$ except that its i-th and j-th components are $k_i - 1$ and $k_j + 1$, respectively. Here the first term of the right hand side is the probability component corresponding to the case that the state $S(\boldsymbol{k})$ is unchanged due to the complete balance of arrival from and departure to outside. The second term corresponds to the case that state gets increased from $S(\boldsymbol{k}) - 1$ to $S(\boldsymbol{k})$, by external arrivals. The third term corresponds to the case that state gets decreased from $S(\boldsymbol{k}) + 1$ to $S(\boldsymbol{k})$, by departure to outside. The fouth term is the case corresponding to the case that the state $S(\boldsymbol{k})$ is unchanged due to just internal emigration-immigration.

For equilibrium of $t \to \infty$, the differential equation can be solved for $\lim_{t\to\infty} P_{\boldsymbol{k}}(t) = p_{\boldsymbol{k}}$ to be

$$p_{\boldsymbol{k}} = \frac{1}{G}\, f(\boldsymbol{k})\, F(S(\boldsymbol{k})), \tag{16}$$

where

$$F(K) = \prod_{S(\boldsymbol{k}=0}^{K-1} \gamma(S(\boldsymbol{k})), \tag{17}$$

$$f(\boldsymbol{k}) = \prod_{i=1}^{N} \prod_{j_i=1}^{k_i} \frac{e_i}{\mu_{j_i}} \qquad \text{and} \tag{18}$$

$$G = \begin{cases} \sum_{K=0}^{\infty} F(K)H(K) & \text{if the sum converges} \\ \infty & \text{otherwise,} \end{cases} \tag{19}$$

with

$$H(K) = \sum_{\boldsymbol{k}:\, S(\boldsymbol{k})=K} f(\boldsymbol{k}). \tag{20}$$

4.2 Departure from the formulation based on Jackson's Network

Though the Jackson's theorem stated above applies in some parts of our setting, it does not in several other points. The applicable point is the network topology. The differences are in the following points:

- The Jackson's network considers the increase or decrease of number of customers (network data size) only by 1 at a time, while we would like the increase or decrease by general integers. This may call for the multiple occurrence [1][2] in Poisson arrivals.
- The Jackson's network do not consider the spatial dynamics of state probability but only termporal, while we would like to consider the data size flow in a spatio-temporal point dynamics.
- We would like to model certain cluster processes cause by the Poisson arrival as well, in order to describe the self-similarity or LRD.

Below we consider a traffic model of UC that takes into account the first three features of the above, to seek for the formulation of traffic model.

For the l-th region, let the aggregated user traffic process be $X^{(l)}(t)$, $t \geq 0$. By $p_n^{(l)}(t)$, we denote $P(X^{(l)}(t) = n)$ for $n \in \mathbb{N}_0$.

First, we would like to describe a differential equation to formulate the $p_n^{(l)}(t)$. Before that, however, we should remark that it will be a better approach to take the state space of the amount of user traffic $X^{(l)}(t)$ rather than the number of connected users as is traditional in classical queueing theory.

The reason for this, besides that basically we would like to estimate the amount of the traffic at a time instant t or as $t \to \infty$, is as follows.

Consider the distribution of amount of user traffics $\varphi \geq 0$ [MB/s]. Then each of the traffics runs over horizontal time axis, as an on-off source [9][17], with φ the vertical axis. Thus φ may be considered to take nonnegative integer values.

It may as well be written with φ real-valued for the sake of theoretical formulation. In such cases, one may wants to divide the vertical axis with a finite number of partition so that the one can apply a traditional birth-death processes in queueing theory in every partitioned classes, with input and output rates depending on the classes. This is awkward, however, since the distribution of user traffics consists of "so many mice and rare elephants" [11][15]. Here the mice mean those users that cause small amount of traffics, while the elephants those that cause extraordinarily large amount of traffics. The distribution of the user traffic thus vary over φ. This may be written well neither by the partitions of φ stated above nor by considering the number of connected users as the vertical axis. Presence of the each traffic size class along time axis is depicted in Figure 5. Therefore, we will consider the differential equation of $p_n^{(l)}(t)$, not of number of users but the amount of aggregated user traffic. Now let λ_ν and μ_ν, $\nu =\in N_0$ be input and output rate, respectively. Then, along with the birth-death state transition in classical queueing theory but with multiple occurrence of points [1], we may write

$$\begin{aligned} p_n^{(l)}(t+h) = p_n(t) \left\{ 1 - \sum_{\nu \in \mathbb{N}} (\lambda_\nu^{(l)} + \mu_\nu^{(l)})\, h \right\} \\ + \sum_{\nu=1}^{n} p_{n-\nu}^{(l)}(t+h)\, \lambda_\nu^{(l)} h \\ + \sum_{\nu \in \mathbb{N}_0} p_{n+\nu}^{(l)}(t+h)\, \mu_\nu^{(l)} h + o(h), \end{aligned} \tag{21}$$

where

$$\begin{cases} \lambda_\nu^{(l)} = \zeta_\nu^{(l)} + \sum_{i \neq l} \eta_\nu^{(i,l)}, \\ \mu_\nu^{(l)} = \gamma_\nu^{(l)} + \sum_{i \neq l} \eta_\nu^{(l,i)}, \end{cases} \tag{22}$$

with $\zeta_\nu^{(l)}$ and $\gamma_\nu^{(l)}$ the birth and death rates in the l-th region, while $\eta_\nu^{(i,l)}$ and $\eta_\nu^{(l,i)}$ the "immigration" rate from i-th to l-th region and "emmigration" rate from l-th to i-th region, respectively. Here the "immigration" and "emmigration" are for the amount of user traffic and not for the number of connected users.

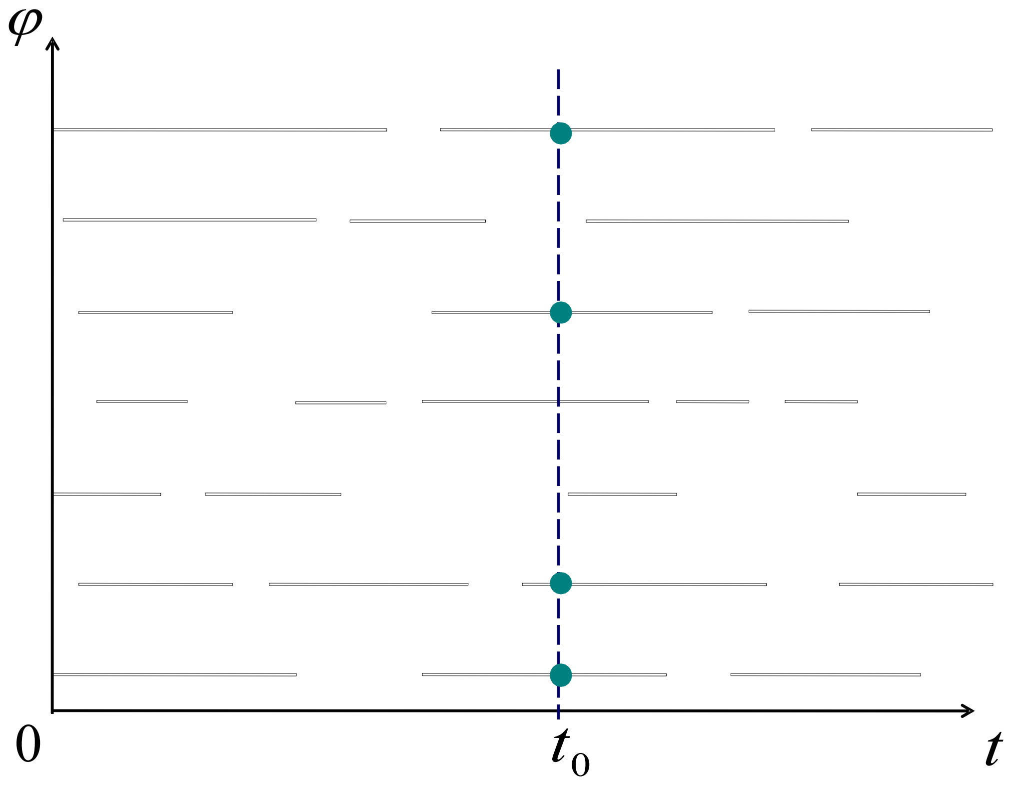

Figure 5: User traffics aggregation at $t = t_0$ over various values of φ

Then (21) yields

$$\frac{dp_n(t)}{dt} = -p_n(t) \sum_{\nu \in \mathbb{N}} (\lambda_\nu^{(l)} + \mu_\nu^{(l)}) + \sum_{\nu=1}^{n} p_{n-\nu}^{(l)}(t+h)\,\lambda_\nu^{(l)} + \sum_{\nu \in \mathbb{N}_0} p_{n+\nu}^{(l)}(t+h)\,\mu_\nu^{(l)}. \tag{23}$$

Solving the equation explicitly may be difficult. Certain another way of characterizing the solution $p_n(t)$ will be desired.

On the other hand, we can write the distribution of aggregated on-off process as follows. Let $I_{\nu,k}$ be the k-th occurrence of the busy period of process for $\varphi = \nu$ and X_ν the random sum of $I_{\nu,k}$ by random upper bound $N_\nu(t) \in \mathbb{N}_0$. Thus we may write

$$X_\nu(t) = \sum_{k=1}^{N_\nu(t)} W_{\nu,k}\, I_{\nu,k}, \tag{24}$$

with the amount load $W_{\nu,k}$ for k-th occurrence for each ν. Then, $X(t) = \sum_\nu X_\nu(t)$ represents the amount of aggregated traffic. Assuming independence of $X_\nu(t)$ over ν, we have

$$\begin{aligned} P(X(t) \le x) &= P\left(\sum_\nu X_\nu(t) \le x \right) \\ &= G_1 * G_2 * \cdots , \end{aligned} \tag{25}$$

where G_ν is the distribution function of $X_\nu(t)$: $G_\nu(t,x) = P(X_\nu(t) \le x)$. Let $\{p_{\nu,k}(t)\} = P(N_\nu(t) = k)$, $\sum_k p_{\nu,k}(t) = 1$ and $F_\nu^{*k}(x)$ the k-th convolution of the distribution $F_\nu(x)$ that t is in some $I_{\nu,k}$:

$$F_\nu(x) = P\left(t \in \sum_{k \in N_0} I_{\nu,k} \right). \tag{26}$$

Then, we have

$$\begin{aligned} G_\nu(t,x) &= P\left(\sum_{k=1}^{N(t)} W_{\nu,k}\, I_{\nu,k} \le x \right) \\ &= \sum_{k \in \mathbb{N}_0} p_{\nu,k}(t)\, F_{\nu,k}^{*k}(x). \end{aligned} \tag{27}$$

Asymptotically as $t \to \infty$, this can be calculated further using a tool of renewal equations [6].

5 Concluding Remarks

In traffic modeling of UC, we have extracted severak problems in spatio-temporal modeling of mobile users and real-time requirement of sending context information and receiving service information. We take into account some characteristics pointed out by several authors after 1990s. While the conventional queueing theoretic models are mainly interested in temporal behavior of the system, we would like to construct a spatio-temporal model for UC, and thus to give a guideline of network design.

The mobile user model in this paper will be a basic concept of the spatio-temporal model. The behavior could be somewhat complicated, but it is due to the spatio-temporal dimension. We are to consider several theoretical problems such as how the movement model is connected to the self-similarity of the aggregated traffic, formulation of the optimal design for a given requirement and whether the limit for large number of aggregation enjoys central limit theorem or not.

Towards the UC traffic model formulation, we can consider several point of view in order to make the model distinct from classical queueing theory. Examples of such point of view may be those listed in the beginning of Section 4. Some of the points are complicated version of basic and classical concepts. Among them, formulation along the last two points may have challenging interdisciplinary interests: the connection of SS and LRD property with spatio-temporal process, and statistical mechanics. We would like to explore the formulation along these directions more.

Finally, it will be necessary to perform a verification of the model through a simulation. For this, one may consider simulating the spatio-temporal dynamics by cell auto-maton. As the cell auto-maton is sometimes used for a microscopic behavior simulation to obtain a macroscopic description of the model, we would like to have an equation, like partial differential equation, that describes the spatio-temporal dynamics.

Acknowledgements: A part of this work was supported by MEXT Grant-in-Aid for Building Strategic Research Infrastructures.

References:

[1] D. R. COX AND V. ISHAM, *Point Processes*, Chapman & Hall, 1980.

[2] D. J. DALEY AND D. VERE-JONES, *An Introduction to the Theory of Point Processes*, Springer, 1998.

[3] M. K. DENKO, Mobility Prediction Schemes in Wireless Ad Hoc Networks, in *Advanced Wired and Wireless Networks*, ed. T. A. Wysocki et al., Springer, 2005.

[4] M. FRANCESCHETTI AND R. MEESTER, *Random Networks for Communication*, Cambridge Univ. Press, 2007.

[5] G. B. FOLLAND, *Introduction to Partial Differential Equations*, Princeton Univ. Press, 1995.

[6] S. KARLIN, *A First Course in Stochastic Processes*, Academic Press, 1975.

[7] J. F. C. KINGMAN, *Poisson Processes*, Oxford Univ. Press, 1993.

[8] L. KLEINROCK, *Queueing Systems, vol.I: Theory*, John Wiley & Sons, 1975.

[9] M. MANDJES, *Large Deviations for Gaussian Queues: Modeling Communication Networks*, Wiley, 2007.

[10] T. MORI, M. UCHIDA AND S. GOTO, Flow Analysis of Internet Traffic: World Wide Web versus Peer-to-Peer, VOL.36, NO.11, Syst. Comp. Japan, 70-81, 2005.

[11] T. Mori, R. Kawahara, S. Naito and S. Goto, On the Characteristics of Internet Traffic Variability: Spikes and Elephants, Proc. IEEE/IPSJ Symp. Appl. Internet (SAINT2004), 99-106, Tokyo, 2004.

[12] S. V. LOWEN, M. C. TEICH, *Fractal-Based Point Processes*, Wiley, 2005.

[13] H. Murakami, The Role of a Mobile Terminal in the Ubiquitous Network, IEEE GLOBECOM 2005.

[14] H. Murakami, R&D of Total Network System Aiming at Realization of Ubiquitous Network Society – From the Results of Ubila Project –, 13th Kyushu International Techno-Fair, 2008.

[15] W. Willinger and V. Paxon, Where Mathematics Meets the Internet, Notices of the AMS, 961-970, 1998.

[16] Y. Takahashi, Application of Queueing Theory to Mobile Communication Systems (in Japanese), Trans. Japan OR Society, 45, 11-20, 2001.

[17] W. Whitt, *Stochastic Process Limits*, Springer, 2002.

[18] S. Kawasaki, M. Higuchi, K. Hirano and H. murakami, A Study on Traffic Modeling in Ubiquitous Cloud, Proceedings of the 8th WSEAS International Conference on Telecommunications and Informatics, 187-193, 2009.

[19] S. M. Torabi and M. A. Samadian, Covering of Problem in Wireless Sensor Networks, Proceedings of the 8th WSEAS International Conference on Telecommunications and Informatics, 88-94, 2009.

[20] S. Rugova and A. Maraj, Traffic Load and Cost Analysis for Different IPTV Architectures, Proceedings of the 8th WSEAS International Conference on Telecommunications and Informatics, 163-168, 2009.

[21] B. D. Ripley and F. P. Kelly, Markov Point Processes, Journal of the London Mathematical Society, s2-15, 188-192, 1977.

[22] R. A. Holley and D. W. Stroock, Nearest neighbor birth and death processes on the real line, Acta Mathematica, 140, 103-154, 1978.

[23] T. M. Liggett, The stochastic evolution of infinite systems of interacting particles, Lecture Notes in Mathematics, 598, 187-248, 1977.