Proceedings of International Conference
on Networking and Multimedia

Performance Study for an ATM Multiplexer with GCRA-Enforced Sources

Jung-Shyr Wu and Fang-Jang Kuo

Department of Electrical Engineering
National Central University, 32054 Chung-Li, Taiwan, Republic of China

Fax: +886-3-4255830

E-mail:jswu@wireless.ee.ncu.edu.tw

Abstract

The Generic Cell Rate Algorithm (GCRA) is recommended by the ATM Forum to perform Usage Parameter Control at the User Network Interface in ATM network where every traffic source is policed by a GCRA enforcer before entering the access node at the edge of the network. In this article, we investigate the characteristics of the model that GCRA-enforced traffic sources are merged together by a multiplexer in the edge node. Traffic parameters concerned include peak cell rate, mean cell rate, CDV tolerance and burst tolerance. Based on the worst output traffic pattern from a dual-stage GCRA enforcer, we obtain the upper bound of the queue length as the function of the GCRA parameters and number of connections.

1. Introduction

Call Admission Control (CAC) and Usage Parameter Control (UPC) are two important steps for congestion control in ATM networks. During the CAC phase, the network user declares the source traffic parameters and the required Quality of Service (QoS), so as for the network to decide whether to accept the connection or not. After a connection is accepted, some UPC scheme must be utilized to monitor and control traffic by detecting whether it conforms to the declared parameters. In the past, the Leaky Bucket (LB) is the most popular one due to its simple algorithm [1,2]. At the edge node of the network, traffic enforced by the UPC scheme is usually multiplexed by a multiplexer before entering the network. Based on the cell loss ratio at the multiplexer, [3] has discussed whether the full rate periodic on/off pattern is the worst pattern. Based on the criteria of average queueing delay at the multiplexer, [4] found out the worst pattern under the assumption that multiple cells can pass through a LB at the same time as long as there are sufficient tokens in the pool.

The ITU-T Recommendation I.371 [5] used the Generic Cell Rate Algorithm (GCRA) to define the traffic parameters Peak Cell Rate (PCR) and Cell Delay Variation Tolerance (CDV Tolerance) of an ATM connection. The ATM Forum [6] applied the algorithm to define Sustainable Cell Rate (SCR, the upper bound of average rate) and Burst Tolerance (BT) so as to facilitate UPC function. There are two pairs of parameters based on GCRA for conformance testing at the User Network Interface (UNI). The PCR (represented by the peak interarrival time I_p) and the CDV Tolerance (denoted as L_p) are tested by GCRA(I_p, L_p). In the same way, the SCR (represented by the mean interarrival time I_s) and the BT are tested by the GCRA(I_s, L_s), where L_s is the sum of BT and the specified CDV Tolerance (L_p) at the UNI. While the parameter BT is conveyed through the expected Maximum Burst Size (MBS) that may be transmitted at peak cell rate according to

$$BT = (MBS-1)(I_s - I_p)$$

So far, the worst traffic pattern after the GCRA enforcer has not been fully discussed in the literature. In this paper, based on the worst output traffic pattern from the GCRA enforcer, we obtain the upper bound on the average queue length as the function of the GCRA parameters and number of connections.

2. Description of system model

The system contains N_s traffic sources. Before they are multiplexed by the multiplexer in the edge node, every source is enforced by a dual-stage GCRA enforcer with a parameter set (I_p, L_p, I_s, L_s). The flow chart of the dual-stage GCRA enforcer is shown in Fig.1. It comprises two single-stage GCRA UPCs. For such a system, we make the following assumptions for the sake of convenient manipulation.

(1). The enforcer only discards non-conforming cells, while it keeps transparent to conforming cells.

(2). The transmission time of the source, the processing time of the enforcer and the propagation delay time are neglected. So, any cell departing from a source would be immediately sent out of the enforcer and be multiplexed to the buffer as long as it is a conforming one.

(3). Any cell found at the input port of the multiplexer is stored in the buffer. If more than one cells arrive at the multiplexer at the same time, the arrival times of these cells at the buffer are treated as the same.

(4). There is sufficient buffer size in the multiplexer such that no cell would be lost.

Due to constant cell size, the service time, which is denoted as $1/\mu$ (or η), is also constant. Such a multiplexer may stand for the one set in front of a switch so as to concentrate user traffic and reduce the required input ports. On the other hand, it may stand for the logical multiplexer at the output port of a switch that collecting cells from different input ports.

For the system to be stable, the total sustainable cell rate must be less than the link capacity, i.e.,

$$\frac{N_s}{I_s} < \mu \ .$$

Besides, if $\frac{N_s}{I_p} \leq \mu$, there would be no bursty level congestion even if all cells arrive at peak cell rate. As a result, for the purpose of nontrivial discussing, we assume that

$$\frac{N_s}{I_p} > \mu \ .$$

3. Most Clumping Pattern (MCP)

In this section, we would like to find out when a cell passes through a dual-stage GCRA and how soon the following cells may pass through it.

As the flow chart shown in Figure 1, if there are some non-conforming cells before the k-th arriving cell, then $k > l$. However, any non-conforming cell would neither pass through the enforcer nor bring a new $TAT_p(l+1)$ or $TAT_s(l+1)$. As a result, we may assume that the sources only send out conforming cells. That is to say, we may simultaneously substitute l with k, and ignore the blocks enclosed by dashed lines in the flow chart and assume that

$$t(k) = \max\{TAT_p(k) - L_p, TAT_s(k) - L_s\}$$

As we can see from Fig. 1,

$$TAT_p(k+1) = \begin{cases} TAT_p(k) + I_p, & \text{if } t(k) \leq TAT_p(k) \\ t(k) + I_p, & \text{if } t(k) > TAT_p(k) \end{cases}$$

As a result, for the k-th cell to minimize $TAT_p(k+1)$, $t(k)$ must be subject to

$$t(k) \leq TAT_p(k) \ .$$

Similarly, for the k-th cell to minimize $TAT_s(k+1)$, $t(k)$ is subject to

$$t(k) \leq TAT_s(k) \ .$$

Besides, the inherent limitation on the transmission line is

$$t(k) = t(k-1) + \eta \ ,$$

where $\eta = \frac{1}{\mu}$ is a unit slot time.

Thus, the earliest time for the k-th cell to be a conforming cell and to minimize the allowable arrival time of the $(k+1)$ -th cell is

$$t_e(k)=\max\{TAT_p(k)-L_p, TAT_s(k)-L_s, t(k-1)+\eta\}$$

Without loss of generality, we may assume the arrival time of the first cell is $t(1)=0$ and set the initial value $TAT_p(1)=TAT_s(1)=0$ such that the first incoming cell is conforming. It results in $TAT_p(2)=I_p$ and $TAT_s(2)=I_s$. Similar as the above derivation, we may let the arrival time of the second cell be

$$t(2)=\max\{I_p-L_p, I_s-L_s, \eta\} \quad (1)$$

Here comes out the problem of determining the maximum of three terms in Eq(1). Similar problem would be encountered when determining the arrival time of the following cells. It involves the comparison among the GCRA parameters and η. It is so complicated due to various conditions. However, we may simplify it by the following procedure.
At first, we define

$$X(\omega)=\omega\eta,$$

$$Y(\omega)=\omega I_p-L_p \quad \text{and}$$

$$Z(\omega)=\omega I_s-L_s, \text{ where } \omega\in R^+.$$

The inherent relation between the mean interarrival time, peak interarrival time and the unit slot time is

$$I_p\geq I_s\geq\eta.$$

They respectively correspond to the slopes of $X(\omega)$, $Y(\omega)$ and $Z(\omega)$ with respect to ω. Three possible relationship among $X(\omega)$, $Y(\omega)$ and $Z(\omega)$ are shown in Figure 2. Initially, $X(\omega)\geq Y(\omega)$ and $X(\omega)\geq Z(\omega)$. We define

$A(I_p, L_p, I_s, L_s)$: the minimum positive integer such that $Y(\omega)\geq X(\omega)$ and $Y(\omega)\geq Z(\omega)$, and

$B(I_p, L_p, I_s, L_s)$: the minimum positive integer such that $Z(\omega)\geq X(\omega)$ and $Z(\omega)\geq Y(\omega)$.

Evidently, B does exist. However, A may not always exist. Now, we divide the relationship between the GCRA parameters and η into two classes. For class 1, A does exist and for class 2, A does not exist. As a result, Figure 2-(a)'s is classified as class 1 while Figure 2-(b)'s is classified as class 2.

Class 1 :

Because both A and B exist, we may describe the discrete characteristics in this class as

$$\begin{cases} X(W)\geq Y(W) \text{ and } X(W)\geq Z(W),\ 0\leq W\leq A-1, \\ Y(W)\geq X(W) \text{ and } Y(W)\geq X(W),\ A\leq W\leq B-1, \\ Z(W)\geq X(W) \text{ and } Z(W)\geq X(W),\ B\leq W, \end{cases}$$

where W may be any positive integer.
So, in this class, we can obtain the result of Eq (1) as $t_1(2)=\eta$ (i.e., $W=1$), where the subscript '1' means class 1. Referring to Fig. 2, we have

$$TAT_p(3)=TAT_p(2)+I_p=2I_p \quad \text{and}$$

$$TAT_s(3)=TAT_s(2)+I_s=2I_s.$$

Then, according to Eq (5), we may let the third cell arrives at the earliest time ($W=2$), i.e.,

$$t_1(3)=t_e(3)=\max\{2I_p-L_p, 2I_s-L_s, 2\eta\}=2\eta$$

Similarly, we may let the i-th cell arrive at the earliest time ($W=i-1$), i.e.,

$$\begin{aligned} t_1 &= \max\{(i-1)I_p-L_p, (i-1)I_s-L_s, (i-1)\eta\} \\ &= (i-1)\eta,\ 1\leq i\leq A,\ i\in N. \end{aligned}$$

Following the same principle we can obtain the earliest arrival time of other cells and express them as,

$$t_1(i)=\begin{cases} (i-1)\eta, & 1\leq i\leq A, \\ (i-1)I_p-L_p, & A+1\leq i\leq B, \\ (i-1)I_s-L_s, & B+1\leq i. \end{cases}$$

If we denote the interarrival times as $\Delta_1(i)=t_1(i+1)-t_1(i)$, then

$$\Delta_1(i)=\begin{cases}\eta, & 1\le i\le A-1,\\ AI_p-L_p-(A-1)\eta, & i=A,\\ I_p, & A+1\le i\le B-1,\\ BI_s-L_s-(B-1)I_p+L_p, & i=B,\\ I_s, & B+1\le i,\end{cases}$$

where $\eta < AI_p-L_p-(A-1)\eta < I_p$ and $I_p < BI_s-L_s-(B-1)I_p+L_p < I_s$ For an arrival pattern with such a sequence of interarrival time, all cells (except the first one) arrive with the shortest allowable time apart from the first one. So, we name such a pattern as Most Clumping Pattern (MCP) of class 1.

Class 2 :

Because only B exists as shown in Figure 2-(b), we may describe the discrete characteristics in this class as

$$\begin{cases}X(W)\ge Y(W) \text{ and } X(W)\ge Z(W), & 0\le W\le B-1,\\ Z(W)\ge X(W) \text{ and } Z(W)\ge Y(W), & B\le W,\end{cases}$$

Similar as the derivation for class 1, we may assign the earliest arrival time of each cell at

$$t_2(i)=\begin{cases}(i-1)\eta, & 1\le i\le B,\\ (i-1)I_s-L_s, & B+1\le i,\end{cases}$$

We similarly define

$\Delta_2(i)=t_2(i+1)-t_2(i)$ and obtain

$$\Delta_2(i)=\begin{cases}\eta, & 1\le i\le B-1,\\ BI_s-L_s-(B-1)\eta, & i=B,\\ I_s, & B+1\le i.\end{cases}$$

For an arrival pattern with such a sequence of interarrival time, all cells (except the first one) arrive with the shortest allowable time. So, we call such a pattern as Most Clumping Pattern (MCP) of class 2.

4. Average queue length

In this section, we shall derive an upper bound for average queue length in the multiplexer buffer. Before that, we shall introduce two lemmas. First, we define some notations as follows.

$\alpha(t)$: It is a funtion representing the number of arriving cells in $[0,t]$. It also stands for the arrival process.

N_b : The number of busy periods.

$\left[t_k^s,t_k^e\right]$, $1\le k\le N_b$: It stand for the range of the k-th busy period, where the superscriptor "s" and "e" represent starting and end points, respectively.

${}_{t_a}^{t_b}L(\alpha(t))=\frac{1}{t_b-t_a}\int_{t_a}^{t_b}(\alpha(t)-\beta(t))dt$: the average queue length corresponding to the arrival process $\alpha(t)$ during the period $[t_a,t_b]$.

Lemma 1.

For a server with constant service times and sufficient buffer size,

$$ {}_0^{t_{end}}L(\alpha(t)) < \max_{k\in N}\left[{}_{t_k^s}^{t_k^e}L(\alpha(t))\right] .$$

proof.

See ref [7].

This lemma allows us to find the upper bound of average queue length only by observing the worst busy period. In the following, we shall investigate how to maximize the average queue length as well as what the quantity is.

Lemma 2.

Assume an arrival process $\alpha(t)$ feeding to a server with constant service times and sufficient buffer size and resulting in a unique busy period $[0,t_{prd}]$. If any of the cell arrival time, except the first one, is shifted forward, and the resulted arrival process is represented as $\alpha^{fw}(t)$, then

$$ {}_0^{t_{prd}}L(\alpha^{fw}(t))\ge {}_0^{t_{prd}}L(\alpha(t)) .$$

Proof.

See ref[7].

Then, according to this lemma, in order to get the upper bound of average queue length in one busy period, we should hasten the arrival time legally for every cell in one busy period.

4.1 Synchronous Clumping Condition (SCC)

We first define the condition that all sources synchronously send out MCP as Synchronous Clumping Condition (SCC). Under this condition, all of the transmitted cells would transparently pass through the dual-stage GCRA. We call the N_s cells that simultaneously arrive at the multiplexer as a bulk (or a batch). For such an aggregated traffic process, we wish to know that at least how many bulks must arrive at the buffer so as to maximize the average queue length. The derivation must include two classes corresponding to that of MCP.

4.1.1 Class 1

To find the maximum of average queue length for class-1 arrival pattern, we shall divide the arrival pattern into three parts according to interarrival time. For convenience, we let the arrival time of the first bulk be $t = 0$ and denote that of the i-th bulk as $t_1(i)$, where the subscript "1" means class 1.

Part I: For the first A bulks, the arrival time of the i-th bulk is

$$t_1(i)=\begin{cases}0, & i=1\\ (i-1)\eta, & 2\le i\le A\end{cases}$$

According to the definition, average queue length during the busy period *(0, t1(j))* is defined as

$$\bar{d}_i=\frac{\sum_{n=1}^{N_s}W(i,n)}{N_s}=\frac{(2i-1)N_s-1}{2\mu}-(i-1)\frac{1}{\mu},\ (1\le i\le N)$$

where the numerator represents the total time all cells have spent in the queue during the busy period *(0, t1(j))*, the denominator represents the busy period and *W1(i,n)* is unfinished work (or waiting time) caused by the *n-th* cell in the *i-th* bulk which can be derived as

$$W(i,n)=\frac{(i-1)N_s+n-1}{\mu}-(i-1)\frac{1}{\mu},\ 1\le i\le N,\ 1\le n\le N_s$$

For convinience, we define

$$\bar{d}_i=\frac{\sum_{n=1}^{N_s}W(i,n)}{N_s}=\frac{(2i-1)N_s-1}{2\mu}-(i-1)\frac{1}{\mu}.$$

Hence the average queue length in the interval *(0, t1(j))* is

$$\overline{W_i}=\frac{\bar{d}_1+\bar{d}_2+\cdots+\bar{d}_i}{i}$$

It follows that $\bar{d}_i-\bar{d}_{i-1}=\frac{N_s-1}{\mu}\ge 0\ (\because N_s\ge 1)$,

so the $\bar{d}_1(j)$ is increasing.

It can be easily checked that

$$\overline{Q}_1(i)>\overline{Q}_1(i-1),\ 2\le i\le A$$

Hence, the maximum of average queue length for Part I is obtained as

$$\overline{Q}_1(A)=\frac{A(N_s-1)}{2}$$

Part II: The bulk arrival times in this part are expressed as

$$t_1(i)=(i-1)I_p-L_p,\ A+1\le i\le B$$

Similarly we have

$$W_1(i,n)=[(i-1)N_s+n-1]\eta-I_p(i-1)+L_p,\ \text{for } A+1\le i\le B$$

and

$$\bar{d}_1(i)=\frac{[(2i-1)N_s-1]\eta-2I_p(i-1)+2L_p}{2}$$

Hence the average queue length in the interval *(0, t1(i))* is

$$\overline{Q}_1(i)=\frac{\sum_{k=1}^{i}\bar{d}_1(k)}{i\eta},\ \text{for } A+1\le i\le B$$

Since we have assumed that $N_s\eta>I_p$, it follows that

$$\bar{d}_1(i)-\bar{d}_1(i-1)=N_s\eta-I_p>0$$

So $\bar{d}_1(i)$ is increasing with i (for $1\le i\le B$). It can be easily proved that

$$\overline{Q}_1(i) > \overline{Q}_1(i-1), \quad 1 \le i \le B$$

Hence, the maximum of average queue length for Part I and Part II occurs at $i = B$ and it is expressed as

$$\overline{Q}_1(B) = \frac{\left\{\left(B^2 N_s - B - A^2 + A\right)\eta - (B-A)\left[(A+B-1)I_p - 2L_p\right]\right\}}{2B\eta}$$

Part III: From the $(B+1)$-th bulk on, the arrival time of the i-th bulk is $t_1(i) = (i-1)I_s - L_s$. To maintain a busy period, the next bulk must arrive before its previous bulks are completely served, thus we have

$$(i-1)I_s - L_s \le (i-1)N_s\eta$$

So, there is a limit on the bulk number in a busy period and the maximal one is

$$\varepsilon(I_p, L_p, I_s, L_s) = \left\lfloor \frac{I_s + L_s - N_s\eta}{I_s - N_s\eta} \right\rfloor$$

Following the derivations similarly as before, we obtain

$$W_1(i,n) = ((i-1)N_s + n - 1)\eta - I_s(i-1) + L_s, \text{ for } B+1 \le i \text{ and } 1 \le n \le N_s$$

$$\overline{d}_1(i) = \frac{\left[(2i-1)N_s - 1\right]\eta - 2I_s(i-1) + 2L_s}{2}$$

, and

$$\overline{Q}_1(i) = \frac{\sum_{k=1}^{i} \overline{d}_1(k)}{i\eta}, \quad \text{for } B+1 \le i \le \varepsilon$$

Since we assumed $N_s\eta < I_s$., it follows that $\overline{d}_1(i) < \overline{d}_1(i-1)$, for $B+1 \le i \le \varepsilon$

Although $\overline{d}_1(i)$ is decreasing with i (for $B+1 \le i \le \varepsilon$), it is not true that $\overline{Q}_1(i)$ is also decreasing for $B+1 \le i \le \varepsilon$. However, it can be easily checked that

$$\begin{cases} \overline{Q}_1(i) \ge \overline{Q}_1(i-1), & \text{if } \overline{d}_1(i) \ge \eta\overline{Q}_1(i-1), \\ \overline{Q}_1(i) < \overline{Q}_1(i-1), & \text{if } \overline{d}_1(i) < \eta\overline{Q}_1(i-1). \end{cases}$$

Summarizing the results analyzed in from Part I to Part III, we make the following statements.

(i). If $\overline{d}_1(B+1) < \overline{Q}_1(B)$, then the maximum of average queue length in a busy period happens at $i = B$ (for $1 \le i \le \varepsilon$) and the maximum is $\tilde{Q}_1 = \overline{Q}_1(B)$.

(ii). On the other hand, if $\overline{d}_1(B+1) \ge \overline{Q}_1(B)$, then the maximum of average queue length happens at a certain i_1^*-th bulk $(B+1 \le i_1^* \le \varepsilon)$, and the maximum is $\tilde{Q}_1 = \overline{Q}_1(i_1^*)$.

The condition that i_1^* exists is that

$$\overline{d}_1(i_1^*) \ge \eta\overline{Q}_1(i_1^* - 1) = \frac{\sum_{k=1}^{i_1^*-1} \overline{d}_1(k)}{i_1^* - 1}, \quad \text{for } B+1 \le i_1^* \le \varepsilon$$

It means that i_1^* must also satisfy the following inequality

$$(\mu I_s - N_s)i_1^{*2} - (\mu I_s - N_s)i_1^* - \left(1 - 2\mu d_1 + B^2 N_s - \mu B^2 I_s + \mu B I_s + 2\mu B L_s\right) \le 0$$

where $d_1 = \sum_{k=1}^{B} \overline{d}_1(k)$ and i_1^* is a positive integer with $B+1 \le i_1^* \le \varepsilon$

In summary, we obtain the maximum of average queue length under the synchronous clumping condition (SCC) as

$$\tilde{Q}_1 = \begin{cases} \overline{Q}_1(B), & \text{if } i_1^* \text{ does not exist,} \\ \overline{Q}_1(i_1^*) & \text{if } i_1^* \text{ exists for } B+1 \le i_1^* \le \varepsilon. \end{cases} \quad (2)$$

4.1.2 Class 2

To find the maximum of average queue length corresponding to class-2 pattern, we shall divide the arrival pattern into two parts for discussion.

Similarly, summarizing the results analyzed in Part I and Part II, we make the following statements.

(i). If $\bar{d}_2(B+1)<\eta\bar{Q}_2(B)$, then the maximum of average queue length in a busy period happens at $i=B$ (for $1\le i\le\varepsilon$) and the maximum is $\tilde{Q}_2=\bar{Q}_2(B)$.

(ii). On the other hand, if $\bar{d}_2(B+1)\ge\eta\bar{Q}_2(B)$, then the maximum of average queue length happens at a certain i_2^*-th bulk $\left(B+1\le i_2^*\le\varepsilon\right)$, and the maximum is $\tilde{Q}_2=\bar{Q}_2\left(i_2^*\right)$

The condition that i_2^* exists is that

$$\bar{d}_2\left(i_2^*\right)\ge\eta\bar{Q}_2\left(i_2^*-1\right)\text{ for } B+1\le i_2^*\le\varepsilon$$

It means that i_2^* must also satisfy the following inequality

$$(\mu I_s-N_s)i_2^{*2}-(\mu I_s-N_s)i_2^*-\left(1-2\mu d_2+B^2N_s-\mu B^2I_s+\mu BI_s+2\mu BL_s\right)\le 0.$$

where $d_2=\dfrac{B^2(N_s-1)\eta}{2}$ and i_2^* is a positive integer which satisfying $B+1\le i_2^*\le\varepsilon$

In summary, we express the maximum of average queue length for class-2 arrival pattern under the SCC as

$$\tilde{Q}_2=\begin{cases}\bar{Q}_2(B) & \text{if } i_2^* \text{ does not exist,}\\ \bar{Q}_2\left(i_2^*\right) & \text{if } i_2^* \text{ exists for } B+1\le i_2^*\le\varepsilon.\end{cases}$$

(3)

4.2 Any condition other than SCC

For any arrival pattern we can show that [7] the maximum of average queue length for the system in Figure 1 is upperly bounded by $\tilde{Q}_1$ as in Eq (2) for class-1 arrival pattern or by $\tilde{Q}_2$ in Eq (3) for class-2 arrival pattern obtained under SCC.

5. Numerical examples and discussion

Example 1.

If we choose I_s=20, $I_p=10$, $L_p=20$, $\eta=1$, $B(\text{maximum burst size})=20$, $N_s=15$

we can get

$$L_s=B_T+L_p=(B-1)(I_s-I_p)+L_p=210$$

cells time.

$$\bar{Q}_1(B)=72.85 \quad \text{cells,}$$

$$A=\left\lfloor 1+\frac{L_p}{I_p-\eta}\right\rfloor=3 \quad ,$$

$\varepsilon(\text{maximum number of bulks in a busy period})=43$

And the arrival pattern belongs to the class-1 pattern, we obtain $0\le i_1^*\le 28$, but i_1^* is also subject to the inequality $21\le i_1^*\le 43$. So the maximum of average queue length happens at i_1^* =28 and the maximum is 80.46 cells.

Example 2.

If we set $I_s=20$, $I_p=10$, $N_s=15$, $B=20$, $L_p=171$, $\eta=1$, the arrival process belongs to class-2 pattern thus we obtain

$$L_s=B_T+L_P=361 \quad ,$$

$\bar{Q}_2(B)=140$ cells time, $\varepsilon=73$. We also have $0\le i_2^*\le 38$,, but i_2^* is also subject to the inequality $21\le i_2^*\le 73$. So the maximum of average queue length happens at $i_2^*=38$ and the maximum is 180.5 cells.

6. Conclusion

In ATM networks, the UPC mechanism plays a key role for congestion control, namely, in order to prevent illegal traffic by malicious users from entering network, each source must be monitored by the UPC mechanism before entering network. Hence, the behavior of aggregated enforced souce at the multiplexer and the relationship with GCRA UPC parameters are worth discussing. In this article, we have found out an upper bound of average queue length for the system in Figure 1. It was obtained when the N_s sources send out cells under synchronous clumping condition (SCC). Although the relation between the GCRA parameters is so complicated, we obtained the result by simply classifying it as two classes only. We also showed that the upper bound under synchronous condition is larger than that for asynchronous condition.

References

[1] G. Niestegge, "The leaky bucket policing method in the ATM network," International Journal of Digital and Analog Communications Systems, vol. 3, pp. 187-197, 1990.

[2] N. Yamanaka, Y. Sato and K. Sato, "Performance limitation of bucket algorithm for usage parameter control and bandwidth allocation methods," IEICE Trans. Commun., vol. E75-B, no. 2, pp. 82-86, Feb. 1992.

[3] B. T. Doshi, "Deterministic rule based traffic descriptors for broadband ISDN: worst case behavior and connection acceptance control," IEEE GLOBECOM'93, pp. 1759-1764.

[4] D. C. Lee, "Effect of Leaky Bucket parameters on the average queue delay: worst case analysis," IEEE INFOCOM'94, pp. 482-489.

[5] ITU-T RECOMMENDATION I.371, Traffic Control and Congestion Control in B-ISDN, Geneva, June 1992.

[6] The ATM Forum 1993, "ATM User-Network Interface Specification, Version 3.0," 1993.

[7] F.J. Kuo, Ph.D thesis, Nat. Central University, 1996 Nov.

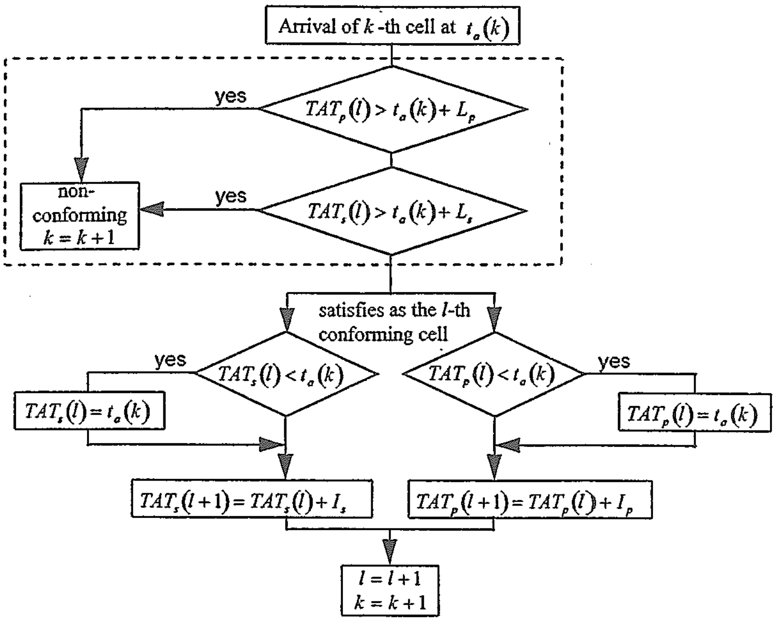

Fig. 1 Dual-Stage GCRA

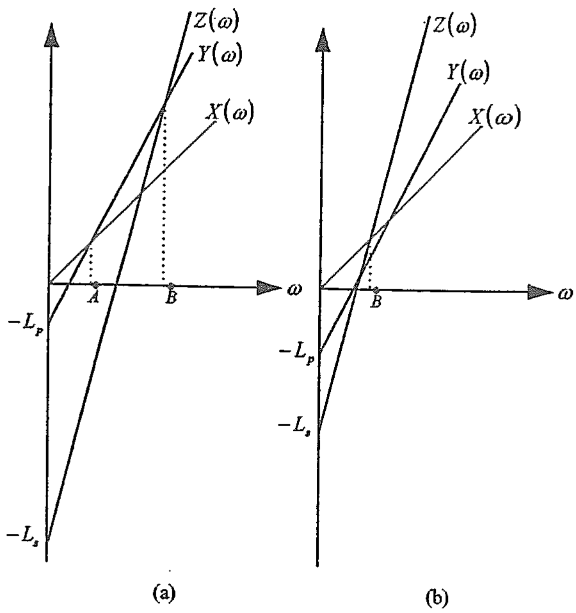

Fig. 2 Two possible relations for $X(\omega)$, $Y(\omega)$ and $Z(\omega)$