

The Timing Capacity of Single-server Queues with Multiple Input and Output Terminals

Xin Liu and R. Srikant

ABSTRACT. We consider an exponential server queue accessed by many different flows. It is assumed that the source and destination information of each flow is available in the packet header. We compute upper and lower bounds on the sum timing capacity of this channel. We discuss the implications of this result for a single flow, where packets can be “colored” to distinguish them, and discuss the role of the scheduling discipline at the server on the amount of uncertainty in the information obtained by an eavesdropper who only observes the packet contents but not the timing between the packets. We also derive asymptotically tight lower and upper bounds on the timing capacity for discrete-time queues with general service-time distributions.

1. Introduction

In [AV96], the authors analyze the timing capacity of continuous-time $M/G/1$ queues. They show that, when the service time distribution is exponential, the timing capacity is e^{-1} nats per average transmission time, which is the lowest among all service time distributions. In other words, if the service rate is μB nats/sec and each packet contains B nats, then the timing capacity of the exponential server is $e^{-1}\mu$ nats/sec. To achieve this timing capacity, packets are sent to the server according to a Poisson process with rate $e^{-1}\mu$.

In this paper, we consider an exponential server queue accessed by many different flows. It is assumed that the source and destination information of each flow is available in the packet header. We study the amount of timing information that can be sent through the system. First, we consider the situation where there are two flows. We are interested in sending information in the inter-packet timing of one flow, while the other flow behaves as interference traffic. We obtain a lower bound on the timing capacity. We then calculate lower and upper bounds on the sum timing capacity of multiple flows. We also study the system in the presence of an *eavesdropper* and the effects of scheduling policies at the server on the amount of additional uncertainty in the information observed by the eavesdropper as compared to the uncertainty at the intended receiver.

Research supported by DARPA through grant F30602-00-2-0542 and AFOSR through URI grant F49620-01-1-0365.

In realistic data networks such as the Internet, information in the packet header is used to distinguish flows from different sources and to different destinations, and thus, to facilitate routing. Given that packets naturally contain information in their headers to distinguish between various flows, the problem we consider in this paper is one of computing the sum timing capacity of many flows accessing a single server. Consider the following example. A protocol reserves B nats in each packet header for address, and thus can distinguish up to $N = e^B$ flows. Suppose that N such flows are queued at a server which processes μ packets per second. Our results show that, for an exponential server queue in continuous-time or for general service-time queues in discrete-time, the sum timing capacity of the N flows is asymptotically close to μB nats per second as $N \rightarrow \infty$. In other words, one can convey almost as much information with timing as one could convey in the traditional manner, i.e., through the bits in the packets.

The N -flow result has an interesting implication for a single flow. Consider a protocol (e.g., IP protocol) that has a fixed number of bits in the header. There are many bits in the header, say equivalent to β nats, that are currently unused and the source can populate these bits in any manner that it wants. We can split a single flow into e^β different sub-flows and use the β nats in the packet header to distinguish between the sub-flows. In this case, the total timing information (i.e., the sum of the information transmitted in the timing between the packets belonging to each sub-flow) of the flow approaches $\beta\mu$ for large β . Thus, a single user can increase its timing capacity by *coloring* its packets, where we use the term coloring to indicate the fact that IDs are used to distinguish between packets of different sub-flows.

One possible use for timing information is to convey information covertly. Typically, for billing/traffic-monitoring reasons, routers (especially those at the edges of a network) record the contents of the packets of some or all of the packets that traverse the network. However, they do not typically measure the timing (inter-arrival times) between packets. From a timing channel point of view, the router can be viewed as an eavesdropper who has access to only packet header information. Thus, the channel model here is similar to the Wyner's wiretap channel [Wyn75]. A natural question is the following: how much timing information can the router extract by just looking at the packet header information of a sequence of packets?

For example, suppose that there are two flows, say RED and GREEN flows, accessing a router and the router records the following sequence of 15 packets going through it: a RED packet, followed by 2 GREEN packets, followed by another RED packet, followed by 10 GREEN packets, and finally a RED packet again. After this packet sequence is recorded at the router, suppose that the network operator obtains information that the RED and GREEN flows have been conveying information through timing and also gains access to the timing code book of both flows. Then, it may be possible to obtain some of the timing information that was conveyed through the queue as follows: it is likely that the timing between the second and third RED packets is larger than the timing between the first and second RED packets since the number of GREEN packets in between the second and third RED packets is larger than the number of GREEN packets in between the first and second RED packets. Thus, it is of interest to quantify the amount of covert information that one can convey using timing when flow identification is available through packet coloring. We partially answer this question by providing bounds on

the additional uncertainty in the information obtained by an eavesdropper who only has access to the packet header information. We also study the effects of service disciplines at the router on the additional uncertainty at the eavesdropper for a single flow with packet coloring.

2. Main Results

We start by studying the timing capacity of a flow which shares a single server with one other flow, which we call the *interference flow*. The existence of interference traffic results in additional randomness (in addition to the randomness due to the service times) in the departure times of packets. We then study the case where multiple input and multiple output flows share the same server. To obtain a bound on the timing capacity in the case of multiple flows, we consider one flow at a time and treat all the other flows together as interference traffic.

2.1. Interference Traffic. Suppose that a sender controls the input process of flow 0 to the server. The server is shared with another flow, denoted by flow I , the interference flow. The sender cannot control or detect the arrivals of the interference packets, and vice-versa. The service discipline is FIFO.

It is interesting to note that, if the service times of flow 0's packets are deterministic, then the sender can achieve infinite timing capacity, regardless the existence of the uncontrollable interference traffic. The reason is that despite the presence of interference traffic, if the total arrival rate is less than the server's service rate, the steady-state probability that a packet belonging to flow 0 sees an empty queue upon arrival is non-zero, and thus, with a positive probability, the decoder can extract timing information without any randomness. One simple strategy to achieve this capacity is to transmit M packets with the same inter-arrival times. The inter-arrival times should be large enough so that the probability of an arrival finding no packets in the queue is positive. If we assume the service time of interference packets is chosen from a continuous distribution, the probability of any two packets have the same *non-zero* waiting time is zero. Given any $\epsilon > 0$, let M be large enough so that the probability that at least three packets from flow 0 incur zero waiting times in the queue is greater than $1 - \epsilon$. Since at least three packets have departed without incurring waiting in the queue (in other words, there are at least two inter-departure times that are exactly equal to the corresponding inter-arrival times), the decoder can measure the inter-arrival time exactly with probability 1, and thus, receive infinite amounts of information.

From now on, we assume that the service times of packets from both flows 0 and I are i.i.d. (independent and identically distributed) random variables with an exponential distribution unless explicitly stated otherwise. The arrival process of flow I is a Poisson process with rate λ_I . The variables A_i and D_i denote the inter-arrival and inter-departure times of the i th packet of flow 0, as shown in Figure 1. In the figure, the solid lines with arrows indicate the arrivals and departures of flow 0's packets and the dashed lines with arrows indicate the arrivals and departures of interference packets. Let \hat{W}_i be the *effective idling time* of the server, i.e., \hat{W}_i is the elapsed time between the departure of the $(i - 1)$ th packet and the arrival of the i th packet of flow 0. In other words, \hat{W}_i is the idling time of the server from

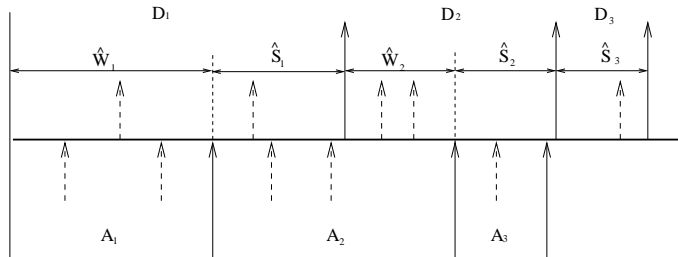


FIGURE 1. Illustration of inter-arrival, inter-departure, and effective service times.

the viewpoint of flow 0. We can write

$$\hat{W}_i = \max\{0, \sum_{j=1}^i A_j - \sum_{j=0}^{i-1} D_j\},$$

where $D_0 = 0$ if the queue is initially empty and D_0 is the system time when packet 0 from flow 0 departs if the queue is initially in equilibrium.

We define \hat{S}_i to be the *effective* service time of the i th packet, which is the time elapsed between the latter of the arrival of the i th packet and the departure of the $(i-1)$ th packet, and the departure of the i th packet. Let n_i be the number of *effective* interference packets for the i th packet of flow 0. To elaborate, if the i th packet arrives before the departure of the $(i-1)$ th packet, n_i is the number of packets from the interference traffic that arrived during A_i . If the i th packet arrives after the departure of the $(i-1)$ th packet, n_i is the number of interference packets in the system when the i th packet arrives. In other words, the i th packet from flow 0 has to wait for n_i interference packets to be served after the departure of the $(i-1)$ th packet. In Figure 1, $n_1 = 1$, $n_2 = 0$, and $n_3 = 1$. In other words, \hat{S}_i is the sum of the service time of the i th packet and that of the n_i interference packets, as shown in Figure 1.

The interference traffic affects the timing information conveyed by flow 0 packets by increasing the randomness of the effective service time. The $\cdot/M/1$ queue studied in [AV96] can be considered as a special case where $n_i \equiv 0$. From the viewpoint of flow 0, packets from the interference traffic that depart when there is no flow 0 packet in the queue do not cause any distortion on the timing, and thus, need not be considered. In Figure 1, the first, third, and fourth interference packets do not have any effect on the departure times of flow 0 packets while the second interference packet does.

PROPOSITION 1. The timing capacity of flow 0 with rate λ_0 , denoted by $C(\lambda_0)$, satisfies

$$(2.1) \quad C(\lambda_0) \geq \lambda_0 \log \frac{\mu - \lambda_I}{\lambda_0}.$$

PROOF. Let the input process of flow 0 be a Poisson process with rate λ_0 and independent of the arrival process of interference traffic. Thus, the input process of the server is Poisson with rate $\lambda_0 + \lambda_I$. Assume that the queue is initially in equilibrium. Burke's theorem for an $M/M/1$ queue [BG87] states that the output process is Poisson with rate $\lambda_0 + \lambda_I$ in steady state. Because the service

discipline is FIFO, each output packet belongs to flow 0 with probability $\lambda_0/(\lambda_0 + \lambda_I)$ independently. Thus, the output process of flow 0 is a Poisson process with rate λ_0 . Let $\bar{D}^n = \{D_0, D_1, \dots, D_n\}$.

We obtain a lower bound on the mutual information between the inter-arrival and inter-departure times of flow 0 in a manner similar to the arguments in [AV96] by noting that the effective service time and the effective idling time of the server play the role of the service time and idling time in the single-flow $M/1$ queue. However, we only obtain a lower bound on the capacity since the effective service times of the packets are not independent. We provide the argument below for completeness:

$$I(A^n; \bar{D}^n) = h(\bar{D}^n) - h(\bar{D}^n | A^n),$$

where

$$\begin{aligned} h(\bar{D}^n | A^n) &= \sum_{i=1}^n h(D_i | A^n, \bar{D}^{i-1}) + h(D_0) \\ &\stackrel{(1)}{=} \sum_{i=1}^n h(D_i | A^n, \bar{D}^{i-1}, \hat{W}_i) + h(D_0) \\ &\leq \sum_{i=1}^n h(\hat{W}_i + \hat{S}_i | \hat{W}_i) + h(D_0) \\ &= \sum_{i=1}^n h(\hat{S}_i | \hat{W}_i) + h(D_0) \\ &\leq \sum_{i=1}^n h(\hat{S}_i) + h(D_0) \\ &\stackrel{(2)}{=} \sum_{i=1}^n \left(\log \frac{1}{\mu - \lambda_I} + 1 \right) + h(D_0), \end{aligned}$$

where (1) holds because \hat{W}_i is a deterministic function of (A^n, \bar{D}^{i-1}) , and the two inequalities hold because conditioning decreases entropy. Note that in the special case where $\lambda_I = 0$, both inequalities hold with equality because $(A^n, \bar{D}^{i-1}) \rightarrow \hat{W}_i \rightarrow D_i$ is a Markov chain and $\hat{S}_i = S_i$ is independent of \hat{W}_i . Last, (2) holds because the \hat{S}_i s are exponentially distributed with mean $1/(\mu - \lambda_I)$ as shown in Lemma 2.1 below. Further, we have

$$h(\bar{D}^n) = \sum_{i=1}^n \left(\log \frac{1}{\lambda_0} + 1 \right) + h(D_0 | D^n),$$

where D_i s are independent and exponentially distributed with mean $1/\lambda_0$ because the output process of flow 0 is Poisson with rate λ_0 . Since

$$\lim_{n \rightarrow \infty} \frac{1}{n} I(D_0; D^n) = 0,$$

we have

$$\lim_{n \rightarrow \infty} \frac{\lambda_0}{n} I(A^n; \bar{D}^n) \geq \lambda_0 \log \frac{\mu - \lambda_I}{\lambda_0}.$$

One way to prove that a lower bound on the mutual information is a lower bound on the capacity is to take the approach in [VH94]; i.e., to show that for

some input process, the inf-information rate is greater than the lower bound. We can do this along the lines of the proof of Theorem 7 in [AV96]. (Because the effective service times, S_i s, are not independent, we cannot directly apply the result of Theorem 7 in [AV96].) To do this, we need to show that $\sum_{i=1}^n D_i/n$ converges in probability to $1/\lambda_0$ and $\sum_{i=1}^n \hat{S}_i/n$ converges in probability to $1/(\mu - \lambda_I)$. The first convergence holds because the output process of flow 0 is a Poisson process with rate λ_0 , and the second one is shown in Lemma 2.1. The other steps of the proof follow as in [AV96]. \square

LEMMA 2.1. *The effective service time, \hat{S}_i , is exponentially distributed with mean $1/(\mu - \lambda_I)$, and $\sum_{i=1}^n \hat{S}_i/n$ converges in probability to $E(\hat{S}_i)$.*

PROOF. Let q_0 be the probability a packet belongs to flow 0, i.e., $q_0 = \lambda_0/(\lambda_0 + \lambda_I)$. Let $\rho = (\lambda_0 + \lambda_I)/\mu$. Because flow 0 and flow I generate independent Poisson arrivals with rate λ_0 and λ_I , a packet belongs to flow 0 with probability q_0 , and it is independent of previous packets. Let $\pi(j)$ be the steady state probability that there are j packets in the queue when the i th packet arrives and $p(k|j)$ be the probability that $n_i = k$ given that there are j packets in the queue when the i th packet arrives. We have

$$\begin{aligned} P(n_i = k) &= \sum_{j=k}^{\infty} \pi(j)p(k|j) \\ &= (1 - \rho)\rho^k(1 - q_0)^k + \sum_{j=k+1}^{\infty} (1 - \rho)\rho^j(1 - q_0)^k q_0 \\ &= \rho^k(1 - q_0)^k(1 - \rho(1 - q_0)) \\ &= \left(\frac{\lambda_I}{\mu}\right)^k \left(1 - \frac{\lambda_I}{\mu}\right), \quad k = 0, 1, 2, \dots \end{aligned}$$

In other words, $n_i + 1$ is a geometrically distributed random variable with mean $\mu/(\mu - \lambda_I)$, and thus

$$E(\hat{S}_i) = \frac{1}{\mu}E(n_i + 1) = \frac{1}{\mu - \lambda_I}.$$

Next, we show that the correlation between the effective service times of packets that are far apart is sufficiently small and thus, one can apply Chebychev's inequality as in the standard proofs of the weak law of large numbers to prove the rest of the lemma. To this end, let

$$g(i) = \{j : \text{the } i\text{th and } j\text{th packets of flow 0 belong to the same busy period}\}.$$

In other words, $g(i)$ is the set of all packets which were served in the same busy period as packet i . Note that \hat{S}_i and \hat{S}_j are independent if i and j do not belong to the same busy period. Let B_i be the duration of the busy period that contains

packet i . Now,

$$\begin{aligned}
E \left[\left(\sum_{i=1}^n \hat{S}_i \right)^2 \right] &= \sum_{i=1}^n E \left[\hat{S}_i \sum_{j \in g(i)} \hat{S}_j \right] + \sum_{i=1}^n E \left[\hat{S}_i \sum_{j \notin g(i)} \hat{S}_j \right] \\
&\leq \sum_{i=1}^n E \left[\hat{S}_i B_i \right] + \sum_{i=1}^n E \left[\hat{S}_i \right] E \left[\sum_{j \notin g(i)} \hat{S}_j \right] \\
&\leq \sum_{i=1}^n E(B_i^2) + \sum_{i=1}^n n E(\hat{S}_i)^2 \\
&= n E(B_1^2) + n^2 \left(\frac{1}{\mu - \lambda_I} \right)^2.
\end{aligned}$$

Thus,

$$\sigma^2(n) = \frac{1}{n^2} \left(E \left[\left(\sum_{i=1}^n \hat{S}_i \right)^2 \right] - E \left[\left(\sum_{i=1}^n \hat{S}_i \right) \right]^2 \right) = \frac{E(B_1^2)}{n}.$$

Because $E(B_1^2)$, the second moment of a busy period of an $M/M/1$ queue, is finite (see e.g., p. 214, (5.142) in [Kle75]) when $\lambda_0 + \lambda_I < \mu$, we have

$$\lim_{n \rightarrow \infty} \sigma^2(n) = 0.$$

By Chebychev's inequality,

$$P \left(\left| \frac{1}{n} \sum_{i=1}^n \hat{S}_i - \frac{1}{\mu - \lambda_I} \right| \geq \epsilon \right) \leq \frac{\sigma^2(n)}{\epsilon^2}.$$

Thus, $n^{-1} \sum_{i=1}^n \hat{S}_i$ converges in probability to $1/(\mu - \lambda_I)$. \square

So far we have provided a lower bound on the timing capacity of a flow in the presence of independent Poisson interference traffic when the service time distribution is exponential. Further, the derivation of the lower bound indicates that the bound is achieved when the input process is a Poisson process, independent of the interference traffic and service times. From the point of view of flow 0, the server behaves *as if* it allocates just enough service to flow I to accommodate its traffic, i.e., the interference traffic gets a service rate of λ_I and flow 0 gets a service rate of $\mu - \lambda_I$. We next use the results of this section to study the timing capacity of multiple flows.

2.2. Multiple Input and Multiple Output Terminals. Let us consider a system where multiple flows share the same server. Information is encoded in the inter-arrival times of packets from the same flow. From the viewpoint of any one flow, all packets from other flows behave as interference traffic. Suppose each flow i , $i = 1, 2, \dots, N$, generates packets according to an independent Poisson process with rate λ_i , where $\sum_i \lambda_i < \mu$. Thus, the output process of each flow is a Poisson process with rate λ_i . By Prop. 1, $R_L(\lambda_1, \dots, \lambda_N)$ is a lower bound on the timing capacity, where

$$R_L(\lambda_1, \dots, \lambda_N) = \sum_{i=1}^N \lambda_i \log \frac{\mu - \sum_{j \neq i} \lambda_j}{\lambda_i}.$$

LEMMA 2.2. *The timing capacity of the N -flow system satisfies*

$$C(\lambda_1, \dots, \lambda_N) \geq R_L(\lambda_1, \dots, \lambda_N) \geq \left(\sum_{i=1}^N \lambda_i \right) \log \frac{\mu}{\sum_{i=1}^N \lambda_i}.$$

Further, the lower bound $R_L(\lambda_1, \dots, \lambda_N)$ is maximized when $\lambda_i = \lambda_j$ for all i and j for a fixed $\sum_{i=1}^N \lambda_i$.

PROOF. First, let $N = 2$. Because one flow treats the other flow as interference, the lower bound on the capacity is

$$R_L(\lambda_1, \lambda_2) = \lambda_1 \log \frac{\mu - \lambda_2}{\lambda_1} + \lambda_2 \log \frac{\mu - \lambda_1}{\lambda_2}.$$

The Hessian matrix of R_L is

$$\nabla^2 R_L = \begin{bmatrix} -\frac{1}{\lambda_1} - \frac{\lambda_2}{(\mu - \lambda_1)^2} & -\frac{1}{\mu - \lambda_1} - \frac{1}{(\mu - \lambda_2)^2} \\ -\frac{1}{\mu - \lambda_1} - \frac{1}{(\mu - \lambda_2)^2} & -\frac{1}{\lambda_2} - \frac{1}{(\mu - \lambda_1)^2} \end{bmatrix}.$$

It is easy to see that the Hessian matrix is negative definite, and thus, R_L is a concave function of (λ_1, λ_2) . Further, R_L is a symmetric function of (λ_1, λ_2) , i.e., $R_L(\lambda_1, \lambda_2) = R_L(\lambda_2, \lambda_1)$. Thus, we have

$$R_L(\lambda_1, \lambda_2) \geq \alpha R_L(\lambda_1 + \lambda_2, 0) + (1 - \alpha) R_L(0, \lambda_1 + \lambda_2) = (\lambda_1 + \lambda_2) \log \frac{\mu}{\lambda_1 + \lambda_2},$$

where $\alpha = \lambda_1/(\lambda_1 + \lambda_2)$. In words, splitting a single flow into two flows increases the timing capacity. By induction, it is easy to see that

$$R_L(\lambda_1, \dots, \lambda_N) \geq \left(\sum_{i=1}^N \lambda_i \right) \log \frac{\mu}{\sum_{i=1}^N \lambda_i}.$$

Further, because R_L is symmetric and concave when $N = 2$, we have

$$R_L(\lambda_1, \lambda_2) \leq R_L\left(\frac{\lambda_1 + \lambda_2}{2}, \frac{\lambda_1 + \lambda_2}{2}\right).$$

Thus, the lower bound is maximized when $\lambda_1 = \lambda_2$. Next, we show it holds for all N . Let $\bar{\lambda} = (\sum_{i=1}^N \lambda_i)/N$. Without loss of generality, we assume that $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_N$. Note that

$$R_L(\lambda_1, \dots, \lambda_N) = R_L(\lambda'_1, \dots, \lambda'_N),$$

where $(\lambda'_1, \dots, \lambda'_N)$ is a permutation of $(\lambda_1, \dots, \lambda_N)$. We have

$$\begin{aligned} R_L(\lambda_1, \lambda_2, \dots, \lambda_{N-1}, \lambda_N) &\leq R_L\left(\frac{\lambda_1 + \lambda_N}{2}, \lambda_2, \dots, \lambda_{N-1}, \frac{\lambda_1 + \lambda_N}{2}\right) \\ &= R_L\left(\lambda_1^{(1)}, \lambda_2^{(1)}, \dots, \lambda_{N-1}^{(1)}, \lambda_N^{(1)}\right) \\ &\leq R_L\left(\frac{\lambda_1^{(1)} + \lambda_N^{(1)}}{2}, \lambda_2^{(1)}, \dots, \lambda_{N-1}^{(1)}, \frac{\lambda_1^{(1)} + \lambda_N^{(1)}}{2}\right) \\ &\vdots \\ &\leq R_L(\bar{\lambda}, \dots, \bar{\lambda}) \\ &= N\bar{\lambda} \log \frac{\mu - (N-1)\bar{\lambda}}{\bar{\lambda}}, \end{aligned}$$

where $(\lambda_1^{(1)}, \lambda_2^{(1)}, \dots, \lambda_{N-1}^{(1)}, \lambda_N^{(1)})$ is a permutation of $((\lambda_1 + \lambda_N)/2, \lambda_2, \dots, \lambda_{N-1}, (\lambda_1 + \lambda_N)/2)$ in decreasing order. In other words, the lower bound is maximized when all flows have the same arrival rates. \square

LEMMA 2.3. *Suppose that there are N flows, $N \geq 2$. Then the overall timing capacity of the N flows, C , satisfies*

$$C \geq \frac{N}{N-1} \mu x^*,$$

where x^* satisfies $x^* e^{x^*} = (N-1)e^{-1}$. To achieve the lower bound, the input processes of the N flows are independent Poisson flows with the same rate

$$\lambda^* = \frac{x^*}{1+x^*} \frac{\mu}{N-1}.$$

PROOF. Let the input process of each flow be an independent Poisson process with rate λ such that $N\lambda \leq \mu$. Each flow considers the other flows as interference traffic. By Prop. 1, we have

$$C \geq N\lambda \log \frac{\mu - (N-1)\lambda}{\lambda}, \quad N\lambda \leq \mu.$$

To find the value of λ that maximizes the lower bound, we take derivative of the right-hand side with respect to λ and set it equal to zero. Let

$$x = \frac{(N-1)\lambda}{\mu - (N-1)\lambda}.$$

The lower bound is maximized when $x^* e^{x^*} = (N-1)e^{-1}$, and we have

$$C \geq \frac{N}{N-1} \mu x^*.$$

The optimal arrival rate of each flow is

$$\lambda^* = \frac{x^*}{1+x^*} \frac{\mu}{N-1}.$$

Because $x e^x$ is an increasing function of x , we have $x^* \leq N-1$. Thus, we verify that $\lambda^* \leq \mu/N$. \square

PROPOSITION 2. Let C be the timing capacity of a server shared by N flows. It satisfies

$$(2.2) \quad (B-1 - \log B)\mu \leq C \leq B\mu,$$

where $B = \log N$, $N \geq 3$.

PROOF. Following Lemma 2.3, we have

$$C \geq \frac{N}{N-1} \mu x^*,$$

where x^* satisfies $x^* e^{x^*} = (N-1)e^{-1}$. Note that $x e^x$ is an increasing function of x . Let

$$x_0 = B-1 - \log B.$$

We have

$$x_0 e^{x_0} = \frac{B-1 - \log B}{B} e^{B-1} \leq \frac{N-1}{N} e^{B-1} = (N-1)e^{-1} = x^* e^{x^*},$$

i.e.,

$$x^* \geq x_0 = B - 1 - \log B.$$

Thus,

$$C \geq \mu(B - 1 - \log B).$$

Next, we show that the maximum timing capacity is upper bounded by μB . Let $X^n = \{X_1, \dots, X_n\}$ be the information sent in the contents of the packets, which is received error-free. Let A^n and D^n be the inter-arrival and inter-departure times of packets passing through the server (which may belong to different flows). The total amount of information passing through the server satisfies:

$$\begin{aligned} I(X^n, D^n; X^n, A^n) &= I(D^n; X^n, A^n) + I(X^n; X^n, A^n | D^n) \\ &\stackrel{(1)}{=} I(D^n; A^n) + I(X^n; X^n), \end{aligned}$$

where (1) holds because X^n contains no additional information regarding D^n other than that in A^n . From [AV96], we know that if $B > 1$ (nat), the maximum information capacity of the channel is μB . Thus,

$$C \leq \sup_{X^n, A^n} I(X^n, D^n; X^n, A^n) \leq \mu B,$$

which concludes the proof. \square

We note that the above theorem shows that the ratio of the upper bound to the lower bound goes to one as B becomes large. Thus, the lower bound on the capacity obtained by considering one flow at a time, and treating the rest of the flows as interference to the first flow, is asymptotically tight.

If the service rate is μ packets per second and each packet contains B nats, then the service rate can be expressed as $\hat{\mu} = \mu B$ nats per second. Suppose there are $N = e^B$ flows in the system, then we have the following result.

COROLLARY 1. The sum timing capacity of the N flows satisfies

$$\left(1 - \frac{1}{B} - \frac{\log B}{B}\right) \hat{\mu} \leq C \leq \hat{\mu}, \text{ where } N \geq 3.$$

Note that if the service rate $\hat{\mu}$ is increased proportional with N , which is often the case in a large system, then the timing capacity *per* flow does not vanish. It is well-known in queueing theory that scaling the capacity with the number of flows has the effect of reducing the waiting time in the system; a fact known as the *statistical multiplexing gain*. However, we note that our result does not exploit the statistical multiplexing gain, i.e., it holds independent of whether or not the system capacity is scaled with the number of flows.

The results in this subsection can be applied to any combination of single sender/receiver, multiple senders/receivers. The upper-bound/lower-bound provided in Prop. 2 bounds the overall timing capacity. To achieve the capacity lower bound, a random encoding strategy for each flow is to assign codewords that are independent realizations of a Poisson process with rate λ^* , and each flow encodes information independently. Each flow decodes information independently by observing the inter-departure times of the flow. There is no collaboration required among different flows. Further, if there exists Poisson cross-traffic with rate λ_I and the service time distribution is exponential with rate μ , then we obtain the capacity bound simply by replacing μ by $\mu - \lambda_I$ in (2.2).

Note that the scheme studied in the paper is different from a multi-user case in the traditional sense, e.g., the case studied in Section 3.2.3 in [Sun99]. In the multi-user case studied in [Sun99], the receiver cannot distinguish between packets from different transmitters. In the current scheme, different flows have different identities and are thus distinguishable. Hence, the result here does not contradict with the result in [Sun99].

2.3. Single Flow with Eavesdropper. In the Internet, the IP protocol requires that each packet contain a certain number of bits in the header. Some of these bits have to be used in a pre-specified manner such as to provide source and destination IP addresses for routing purposes. However, there are also additional bits in the header that are typically unused, but which the source can populate in any manner that it wishes. Let B denote the number of header nats that are not under the source's control and β denote the number of header nats that the flow can use in any way that it wants. Further, suppose that the flow is accessing an exponential-server queue with a service rate of $\hat{\mu}$ nats per second. From [AV96], the maximum timing capacity of this system is $e^{-1}\hat{\mu}/(B + \beta)$ nats per second. A natural question to ask is whether we can increase the timing capacity of the user by appropriately using the β nats. Suppose that we split the single flow into multiple e^β sub-flows and use the β nats to assign a unique id, which we refer to as the *color*, to each sub-flow. Packets with the same color belong to the same sub-flow, and information is encoded in the inter-arrival times between packets of the same color. Thus, by Prop. 2, we have the following result.

COROLLARY 2. If $\beta \geq \log 3$, the timing capacity of a single flow with packet size $B + \beta$ nats satisfies

$$\left(1 - \frac{1}{\beta} - \frac{\log \beta}{\beta}\right) \frac{\hat{\mu}\beta}{\beta + B} \leq C \leq \frac{\hat{\mu}\beta}{\beta + B}.$$

From the lower bound in the above corollary, it is clear that, for sufficiently large β , the timing capacity can be increased by splitting a single flow into multiple sub-flows using packet coloring. It should be clear from the proof of Proposition 2 that the receiver needs to only record the timing between successive packets belonging to a sub-flow to achieve the lower bound. The coding and decoding scheme used to obtain the lower bound does not use the timing information between packets belonging to different sub-flows. However, this information can be useful in the presence of an eavesdropper as we will show in the rest of this subsection.

Consider a system where a sender transmits a private message to a receiver, while an eavesdropper observes the packets passing through the server. The eavesdropper records the sequence and contents of the packets, but not their timings. As noted in Section 1, just by looking at the sequence of packet colors, the eavesdropper obtains some information about the timing between the packets. In what follows, we quantify the amount of the additional uncertainty in the information about the private message received by the eavesdropper (as compared to the uncertainty at the intended receiver). Further, if the eavesdropper is located at the output of the queue, we show that this additional uncertainty is affected by the service discipline used in the queue. Specifically, we show that, under certain service disciplines, packet coloring used to distinguish between sub-flows can increase the amount of additional uncertainty in the private message at the eavesdropper.

Let us consider the case where there are two (sub)flows accessing an exponential server which serves at rate μ packets/sec. Let λ_1 and λ_2 be the data rates (packets-per-second) of flow 1 and flow 2, respectively. Suppose that flow i uses its first n_i packets to transmit timing information, and let $n = n_1 + n_2$. Let A^{n_1} denote the sequence of inter-arrival times of flow 1, T^{n_2} the sequence of inter-arrival times of flow 2. Let D^n be the sequence of inter-departure times of the aggregate flow, S^n the service times of the packets (S_i is the service time of the i th departure), and N^n the sequence of colors of the departing packets, i.e, if $N_k = i$, then the k th departure packet belongs to flow i .

Let

$$R_e = \lim_{n \rightarrow \infty} \frac{\lambda_1 + \lambda_2}{n} I(A^{n_1}, T^{n_2}; N^n)$$

be the information rate obtained at the eavesdropper, and

$$R_r = \lim_{n \rightarrow \infty} \frac{\lambda_1 + \lambda_2}{n} I(A^{n_1}, T^{n_2}; N^n, D^n)$$

be the total information rate observed at the receiver, where it is assumed that $\lim_{n \rightarrow \infty} n_1/n = \lambda_1/(\lambda_1 + \lambda_2)$. Then,

$$R_c = R_r - R_e$$

is the additional uncertainty rate at the eavesdropper as compared to the intended receiver. In this section, we quantify this uncertainty as a function of the service discipline.

We begin by considering the case where the service discipline is FIFO. We note that in this subsection we only consider exponential service time distribution. Let the input processes of flow 1 and flow 2 be Poisson processes with rate λ_1 and λ_2 , respectively.

PROPOSITION 3. When the service discipline is FIFO, the maximum additional uncertainty rate, R_c , at the eavesdropper is equal to the timing capacity of a single flow with rate $\lambda_1 + \lambda_2$.

PROOF. Assume the queue is initially in equilibrium. Packet 0 is sent into the queue as a probing packet and D_0 is the time packet 0 departs (see [AV96]). Recall that $\bar{D}^n = \{D_0, D_1, \dots, D_n\}$. The first n_1 (n_2) in flow 1 (2) are used to convey timing information for flow 1 (2). At the eavesdropper, we have

$$I(A^{n_1}, T^{n_2}; N^n) = H(N^n) - H(N^n | A^{n_1}, T^{n_2}).$$

At the receiver, we have

$$\begin{aligned} I(A^{n_1}, T^{n_2}; N^n, \bar{D}^n) &= I(A^{n_1}, T^{n_2}; \bar{D}^n) + I(A^{n_1}, T^{n_2}; N^n | \bar{D}^n) \\ &= I(A^{n_1}, T^{n_2}; \bar{D}^n) + H(N^n | \bar{D}^n) - H(N^n | A^{n_1}, T^{n_2}, \bar{D}^n) \\ &= I(A^{n_1}, T^{n_2}; \bar{D}^n) + H(N^n) - H(N^n | A^{n_1}, T^{n_2}, \bar{D}^n), \end{aligned}$$

where $H(N^n) = H(N^n | \bar{D}^n)$ because the packet color sequence N^n is independent of the inter-departure sequence \bar{D}^n . Thus, the additional uncertainty at the eavesdropper is

$$I(A^{n_1}, T^{n_2}; \bar{D}^n) + H(N^n | A^{n_1}, T^{n_2}) - H(N^n | A^{n_1}, T^{n_2}, \bar{D}^n).$$

Next, we show that

$$\lim_{n \rightarrow \infty} \frac{\lambda_1 + \lambda_2}{n} I(A^{n_1}, T^{n_2}; \bar{D}^n) = C,$$

where C is the timing capacity of an exponential queue with rate $(\lambda_1 + \lambda_2)$. We have

$$\begin{aligned}
 I(A^{n_1}, T^{n_2}; \bar{D}^n) &= h(\bar{D}^n) - h(\bar{D}^n | A^{n_1}, T^{n_2}) \\
 &= -I(D_0; D^n) + \sum_{i=1}^n h(D_i) - h(D_i | A^{n_1}, T^{n_2}, \bar{D}^{i-1}) \\
 (2.3) \quad &= -I(D_0; D^n) + \sum_{i=1}^n h(D_i) - h(D_i | A^{n_1}, T^{n_2}, \bar{D}^{i-1}, 1_{Z_i}),
 \end{aligned}$$

where 1_{Z_i} is the indicator function of the event Z_i . Event Z_i is the event that

$$\sum_{j=0}^{i-1} D_j \leq \min \left(\sum_{j=1}^{n_1} A_j, \sum_{j=1}^{n_2} T_j \right),$$

which is determined by $(A^{n_1}, T^{n_2}, \bar{D}^{i-1})$. In words, Z_i is the event that the $(i-1)$ th packet departs before either n_1 packets arrive from flow 1 or n_2 packet arrives from flow 2. When Z_i happens, W_i , the idling time of the server between the departure of the $(i-1)$ th packet and the arrival of the i th packet, is determined by $(A^{n_1}, T^{n_2}, \bar{D}^{i-1})$. Thus,

$$\begin{aligned}
 h(D_i | A^{n_1}, T^{n_2}, \bar{D}^{i-1}, 1_{Z_i}) &= h(S_i + W_i | A^{n_1}, T^{n_2}, \bar{D}^{i-1}, 1_{Z_i} = 1, W_i) P(Z_i) \\
 &\quad + h(D_i | A^{n_1}, T^{n_2}, \bar{D}^{i-1}, 1_{Z_i} = 0) (1 - P(Z_i)).
 \end{aligned}$$

Thus,

$$h(S_i) P(Z_i) \leq h(D_i | A^{n_1}, T^{n_2}, \bar{D}^{i-1}) \leq h(S_i) P(Z_i) + h(D_i) (1 - P(Z_i)).$$

Substitute into (2.3), we have

$$\begin{aligned}
 (2.4) \quad &-I(D_0; D^n) + \sum_{i=1}^n P(Z_i) (h(D_i) - h(S_i)) \leq I(A^{n_1}, T^{n_2}; \bar{D}^n) \\
 &\leq -I(D_0; D^n) + \sum_{i=1}^n P(Z_i) (h(D_i) - h(S_i)) + (1 - P(Z_i)) h(D_i).
 \end{aligned}$$

Next, we show that $n^{-1} \sum_{i=1}^n (1 - P(Z_i))$ converges to 0. Note that $\sum_{i=1}^n (1 - P(Z_i))$ is the average number of packets (among the n packets) departs after $\min(\sum_{j=1}^{n_1} A_j, \sum_{j=1}^{n_2} T_j)$. Thus, we have

$$\sum_{i=1}^n (1 - P(Z_i)) \leq L_1 + L_2 + L_3,$$

where L_1 is the average number of packets arrived during $\min(\sum_{j=1}^{n_1} A_j, \sum_{j=1}^{n_2} T_j)$ and $\max(\sum_{j=1}^{n_1} A_j, \sum_{j=1}^{n_2} T_j)$, and L_2 and L_3 are the average numbers of packets in the busy periods that contain the n_1 th packet of flow 1 and the n_2 th packet of flow 2, respectively. Note that L_2 and L_3 are finite. Further, because

$$\lim_{n_1 \rightarrow \infty} \frac{\lambda_1 \sum_{j=1}^{n_1} A_j}{n_1} = \lim_{n_2 \rightarrow \infty} \frac{\lambda_2 \sum_{j=1}^{n_2} T_j}{n_2} = 1,$$

we have

$$\lim_{n \rightarrow \infty} \frac{\sum_{j=1}^{n_2} T_j}{\sum_{j=1}^{n_1} A_j} = \lim_{n \rightarrow \infty} \frac{\min(\sum_{j=1}^{n_1} A_j, \sum_{j=1}^{n_2} T_j)}{\max(\sum_{j=1}^{n_1} A_j, \sum_{j=1}^{n_2} T_j)} = 1,$$

and thus

$$\lim_{n \rightarrow \infty} \frac{L_1}{n} = 0.$$

Substitute it into (2.4). Since the output process is Poisson, $h(D_i)$ is maximized. We have

$$\lim_{n \rightarrow \infty} \frac{\lambda_1 + \lambda_2}{n} I(A^{n_1}, T^{n_2}; \bar{D}^n) = (\lambda_1 + \lambda_2) \log \frac{\mu}{\lambda_1 + \lambda_2} = C,$$

because S_i and D_i are exponentially distributed with mean $1/\mu$ and $1/(\lambda_1 + \lambda_2)$, respectively,

Last, we need to show that

$$\lim_{n \rightarrow \infty} \frac{1}{n} (H(N^n | A^{n_1}, T^{n_2}) - H(N^n | A^{n_1}, T^{n_2}, \bar{D}^n)) = 0,$$

which is enough to show that

$$\lim_{n \rightarrow \infty} \frac{1}{n} H(N^n | A^{n_1}, T^{n_2}) = 0.$$

Let Y_i is the event that the i th packet arrives before $\min(\sum_{j=1}^{n_1} A_j, \sum_{j=1}^{n_2} T_j)$. If event Y_i happens, N_i is determined by $(A^{n_1}, T^{n_2}, N^{i-1})$ when the service discipline is FIFO. Thus, we have

$$\begin{aligned} H(N^n | A^{n_1}, T^{n_2}) &= \sum_{i=1}^n H(N_i | A^{n_1}, T^{n_2}, N^{i-1}, 1_{Y_i}) \\ &= \sum_{i=1}^n H(N_i | A^{n_1}, T^{n_2}, N^{i-1}, 1_{Y_i} = 0) (1 - P(Y_i)) \\ &\leq \sum_{i=1}^n (1 - P(Y_i)). \end{aligned}$$

Similarly, we can show that $n^{-1} \sum_{i=1}^n (1 - P(Y_i))$ converges to 0. Thus, the maximum additional uncertainty rate of two flows at the eavesdropper equals the timing capacity of a single flow with rate $\lambda_1 + \lambda_2$. \square

By coloring and distinguishing between sub-flows, one can increase the timing capacity of a single flow. However, the previous proposition shows that there is no advantage to be gained in this manner if one is interested in transmitting covert information through timing.

Next, we consider the case where the eavesdropper is located at the output of the queue, i.e., it observes the sequence of departing packets, and a random service discipline is employed by the server, i.e., the server randomly picks a packet in queue to serve. The following proposition shows that the difference in the uncertainty between the intended receiver and the eavesdropper can be increased.

PROPOSITION 4. Suppose that the service discipline is random service. The additional uncertainty rate at the eavesdropper, R_e , can be made strictly larger than the timing capacity of a single flow with rate $\lambda_1 + \lambda_2$.

PROOF. Assume the queue is initially in equilibrium. The input processes from flows 1 and 2 are independent Poisson processes with rate λ_1 and λ_2 , respectively. The first n_1 (n_2) in flow 1 (2) are used to convey timing information for flow 1 (2). Let $\tilde{A} = \{\hat{A}_1, \hat{A}_2, \dots, \hat{A}_n, \hat{A}_{n+1}, \dots\}$ be the inter-arrival times of the aggregate flow,

and $\tilde{N} = \{\hat{N}_1, \hat{N}_2, \dots, \hat{N}_n, \hat{N}_{n+1}, \dots\}$ be the ID sequence of the aggregate flow on the arrival. Thus, \hat{A}_i 's are i.i.d. random variables with an exponential distribution, and \hat{N}_i 's are i.i.d. random variables with a Bernoulli distribution. At the receiver, we have

$$I(A^{n_1}, T^{n_2}; N^n, \bar{D}^n) = I(\tilde{A}, \tilde{N}; N^n, \bar{D}^n) - I(\tilde{A}, \tilde{N}; N^n, \bar{D}^n | A^{n_1}, T^{n_2}).$$

We have

$$\begin{aligned} I(\tilde{A}, \tilde{N}; N^n, \bar{D}^n) &= I(\tilde{A}, \tilde{N}; \bar{D}^n) + I(\tilde{A}, \tilde{N}; N^n | \bar{D}^n) \\ &= I(\tilde{A}, \tilde{N}; \bar{D}^n) + H(N^n | \bar{D}^n) - H(N^n | \tilde{A}, \tilde{N}, \bar{D}^n). \end{aligned}$$

Next, we show that $H(N^n | \bar{D}^n) = H(N^n)$ for the random policy. It is equivalent to saying that $I(N^n; \bar{D}^n) = 0$. We have

$$I(N^n; \bar{D}^n) \leq I(N^n; \hat{A}^n, S^n, D_0) = I(N^n; \hat{A}^n, S^n, D_0).$$

The equality holds because \bar{D}^n can be determined by (\hat{A}^n, S^n, D_0) . Now, since (\hat{A}^n, S^n, D_0) cannot give us any information about the identity of a particular packet, $I(N^n; \hat{A}^n, S^n, D_0) = 0$. Thus, $H(N^n | \bar{D}^n) = H(N^n)$. In summary,

$$\begin{aligned} I(A^{n_1}, T^{n_2}; N^n, \bar{D}^n) &= I(\tilde{A}, \tilde{N}; \bar{D}^n) + H(N^n) - H(N^n | \tilde{A}, \tilde{N}, \bar{D}^n) \\ &\quad - I(\tilde{A}, \tilde{N}; N^n, \bar{D}^n | A^{n_1}, T^{n_2}), \end{aligned}$$

where the information rate from the packet header sequence at the eavesdropper is

$$I(N^n; A^{n_1}, T^{n_2}) = H(N^n) - H(N^n | A^{n_1}, T^{n_2}) \leq H(N^n) - H(N^n | \tilde{A}, \tilde{N}).$$

Thus, the additional uncertainty at the eavesdropper is given by

$$\begin{aligned} &I(\tilde{A}, \tilde{N}; \bar{D}^n) + H(N^n | A^{n_1}, T^{n_2}) - H(N^n | \tilde{A}, \tilde{N}, \bar{D}^n) - I(\tilde{A}, \tilde{N}; N^n, \bar{D}^n | A^{n_1}, T^{n_2}) \\ &\geq I(\hat{A}^n; \bar{D}^n) + H(N^n | \tilde{A}, \tilde{N}) - H(N^n | \tilde{A}, \tilde{N}, \bar{D}^n) - I(\tilde{A}, \tilde{N}; N^n, \bar{D}^n | A^{n_1}, T^{n_2}). \end{aligned}$$

Note that $I(\hat{A}^n; \bar{D}^n)$ is the mutual information between the inter-arrivals and inter-departures of a single flow where the input process is Poisson with rate $\lambda_1 + \lambda_2$, and thus

$$\lim_{n \rightarrow \infty} \frac{\lambda_1 + \lambda_2}{n} I(\hat{A}^n; \bar{D}^n) = C,$$

where C is the timing capacity of a single flow with rate $\lambda_1 + \lambda_2$.

Next, we show that

$$(2.5) \quad \lim_{n \rightarrow \infty} \frac{1}{n} \left(H(N^n | \tilde{A}, \tilde{N}) - H(N^n | \tilde{A}, \tilde{N}, \bar{D}^n) \right) > 0,$$

and

$$(2.6) \quad \lim_{n \rightarrow \infty} \frac{1}{n} I(\tilde{A}, \tilde{N}; N^n, \bar{D}^n | A^{n_1}, T^{n_2}) = 0.$$

We have

$$H(N^n | \tilde{A}, \tilde{N}) - H(N^n | \tilde{A}, \tilde{N}, \bar{D}^n) = \sum_{i=1}^n I(N_i; \bar{D}^n | \tilde{A}, \tilde{N}, N^{i-1}).$$

For $i \geq 2$, we have

$$\begin{aligned} I(N_i; \bar{D}^n | \tilde{A}, \tilde{N}, N^{i-1}) &\geq I(N_i; D_{i-1} | \tilde{A}, \tilde{N}, \bar{D}^{i-2}, N^{i-1}) \\ (2.7) \quad &= I(N_i; D_{i-1} | \tilde{A}, \tilde{N}, \bar{D}^{i-2}, N^{i-1}, M_1(i-1), M_2(i-1)), \end{aligned}$$

where $M_1(i-1)$ ($M_2(i-1)$) is the number of flow 1 (2) packets in the queue right after the $(i-1)$ th packet begins its service. Note that $M_1(i-1)$ and $M_2(i-1)$ are determined by $(\tilde{A}, \tilde{N}, \tilde{D}^{i-2}, N^{i-1})$. Thus,

$$(2.7) \geq I(N_i; D_{i-1} | \tilde{A}, \tilde{N}, \tilde{D}^{i-2}, N^{i-1}, M_1(i-1) = 1, M_2(i-1) = 1) \\ \times P(M_1(i-1) = 1, M_2(i-1) = 1).$$

Note that

$$I(N_i; D_{i-1} | \tilde{A}, \tilde{N}, \tilde{D}^{i-2}, N^{i-1}, M_1(i-1) = 1, M_2(i-1) = 1) \\ = I(N_i; S_{i-1} + W_{i-1} | \tilde{A}, \tilde{N}, \tilde{D}^{i-2}, N^{i-1}, M_1(i-1) = 1, M_2(i-1) = 1, W_{i-1} = 0) \\ = I(N_i; S_{i-1} | \tilde{A}, \tilde{N}, \tilde{D}^{i-2}, N^{i-1}, M_1(i-1) = 1, M_2(i-1) = 1, W_{i-1} = 0),$$

where W_{i-1} is the idling time of the server between the departure of the $(i-2)$ th packet and the arrival of the $(i-1)$ th packet. Because the input process is Poisson, $(M_1(i-1) = 1, M_2(i-1) = 1)$ implies that $(W_{i-1} = 0)$. Let t be the time when the $(i-1)$ th packet begins its service. Note that conditioning on $(M_1(i-1) = 1, M_2(i-1) = 1)$, N_i only depends on events that happen after time t . Let

$$\tilde{A}' = \{\hat{A}', \hat{A}_{i+3}, \hat{A}_{i+4}, \dots\} \\ \tilde{N}' = \{\hat{N}_{i+2}, \hat{N}_{i+3}, \hat{N}_{i+4}, \dots\}$$

be the inter-arrival times and ID sequence of packets arrived after time t , where \hat{A}' is a random variable denoting the time elapsed between t and the first packet arrives after t . Note that \hat{A}' is exponential with mean $1/(\lambda_1 + \lambda_2)$. Thus, \tilde{A}' and \tilde{A} have the same distribution, and so do \tilde{N}' and \tilde{N} . Note that \tilde{A}' and \tilde{N}' are independent of $(M_1(i-1) = 1, M_2(i-1) = 1)$. Then, we have

$$I(N_i; D_{i-1} | \tilde{A}, \tilde{N}, \tilde{D}^{i-2}, N^{i-1}, M_1(i-1) = 1, M_2(i-1) = 1) \\ = I(N_i; S_{i-1} | \tilde{A}', \tilde{N}', M_1(i-1) = 1, M_2(i-1) = 1) \\ = \int E \left(\log \frac{p(N_i | s_{i-1}, \tilde{a}, \tilde{n}, M_1(i-1) = 1, M_2(i-1) = 1)}{p(N_i | \tilde{a}, \tilde{n}, M_1(i-1) = 1, M_2(i-1) = 1)} \right) dF(s_{i-1}, \tilde{a}, \tilde{n}).$$

Now, we show that $I(N_i; S_{i-1} | \tilde{A}', \tilde{N}', M_1(i-1) = 1, M_2(i-1) = 1)$ is positive. Consider two events that occur with positive probabilities: 1) during the service time of the $(i-1)$ th packet, no new packet arrives. 2) during the service time of the $(i-1)$ th packet, one new packet arrives. In these two cases, $p(N_i = k | s_{i-1}, \tilde{a}, \tilde{n}, M_1(i-1) = 1, M_2(i-1) = 1)$ is different. Thus, $p(N_i = k | s_{i-1}, \tilde{a}, \tilde{n}, M_1(i-1) = 1, M_2(i-1) = 1) \neq p(N_i = k | \tilde{a}, \tilde{n}, M_1(i-1) = 1, M_2(i-1) = 1)$. Thus, $c_1 =: I(N_i; S_{i-1} | \tilde{A}', \tilde{N}', M_1(i-1) = 1, M_2(i-1) = 1) > 0$. Further, because $P(M_1(i-1) = 1, M_2(i-1) = 1) = c_2 > 0$ and is independent of i in steady state, we have

$$I(N_i; \tilde{D}^n | \tilde{A}, \tilde{N}, N^{i-1}) \geq c_1 c_2 > 0,$$

and thus (2.5) holds. Intuitively, knowing the departure times help reduce the randomness caused by the random policy at the router conditioning on the knowledge of arrivals.

Next, we show (2.6). We have

$$\begin{aligned}
& I(\bar{D}^n, N^n; \tilde{A}, \tilde{N} | A^{n_1}, T^{n_2}) \\
&= I(D_0; \tilde{A}, \tilde{N} | A^{n_1}, T^{n_2}) + \sum_{i=1}^n I(D_i, N_i; \tilde{A}, \tilde{N} | A^{n_1}, T^{n_2}, \bar{D}^{i-1}, N^{i-1}) \\
&= \sum_{i=1}^n I(D_i, N_i; \tilde{A}, \tilde{N} | A^{n_1}, T^{n_2}, \bar{D}^{i-1}, N^{i-1}, 1_{Z_i}),
\end{aligned}$$

where 1_{Z_i} is the indicator function of the event Z_i . Recall that Z_i is the event that the $(i-1)$ th packet departs before either n_1 packets arrive from flow 1 or n_2 packets arrive from flow 2. When Z_i happens, (\tilde{A}, \tilde{N}) does not contain additional information about (D_i, N_i) conditioning on $(A^{n_1}, T^{n_2}, \bar{D}^{i-1}, N^{i-1})$. Thus,

$$\begin{aligned}
& I(D_i, N_i; \tilde{A}, \tilde{N} | A^{n_1}, T^{n_2}, \bar{D}^{i-1}, N^{i-1}, 1_{Z_i}) \\
&= I(D_i, N_i; \tilde{A}, \tilde{N} | A^{n_1}, T^{n_2}, \bar{D}^{i-1}, N^{i-1}, 1_{Z_i} = 0)(1 - P(Z_i)) \\
&\leq \left(I(D_i; \tilde{A}, \tilde{N} | A^{n_1}, T^{n_2}, \bar{D}^{i-1}, N^{i-1}, 1_{Z_i} = 0) + H(N_i) \right) (1 - P(Z_i)) \\
&\leq \left(h(D_i) - h(D_i | \tilde{A}, \tilde{N}, \bar{D}^{i-1}) + H(N_i) \right) (1 - P(Z_i)) \\
&= (h(D_i) - h(S_i) + 1) (1 - P(Z_i)) \\
&= \left(\log \frac{\mu}{\lambda_1 + \lambda_2} + 1 \right) (1 - P(Z_i)).
\end{aligned}$$

Since $\sum_{i=1}^n n^{-1}(1 - P(Z_i))$ converges to 0 as shown earlier, we have shown (2.6), and thus complete the proof. \square

In summary, in the presence of two (sub)flows, if the service discipline is FIFO or if the eavesdropper is located before the server (i.e., observes the packet headers of the arriving packets), then the maximum additional uncertainty at the eavesdropper (as compared to the uncertainty at the intended receiver) equals the timing capacity of a single flow whose arrival rate is the sum of the arrival rates of the two sub-flows. If the service discipline is random, then this additional uncertainty increases. The intuition is that the eavesdropper does not observe or record departure times. The intended receivers' knowledge of the the departure times reduces the randomness introduced by the random service discipline conditioned on the arrival times. Thus, the additional uncertainty at the eavesdropper increases.

Note that it has been shown that for a discrete memoryless channel, if $R_c \leq I(X; Y) - I(X; Z)$, where X is the input random variable, Y and Z are the output random variables observed at the receiver and the eavesdropper respectively, then R_c can be achieved with perfect secrecy [Wyn75, CK78]. Perfect secrecy means that the transmitter can reliably send information to the receiver while the information obtained by the eavesdropper is arbitrarily small [CK78]. We are not aware of any coding theorems that show this for our model. Thus, an open problem is to check whether the transmitter and the receiver can communicate at rate R_c with perfect secrecy in our timing channel.

2.4. Discrete Case. In [BA98], the authors study the timing capacity of discrete queues. Their first model assumes single packet arrivals and departures in a time slot and independent service times. In such a model, the geometric distribution plays a role similar to exponential distribution in the continuous case.

In other words, when the service time is geometrically distributed, the server has the least timing capacity among all distributions with the same average service time, and to achieve the capacity, the input process is Bernoulli process.

In this section, we consider a discrete single server queue shared by many flows. We assume single packet from each flow arrives in a time slot and single departure in a time slot. The service discipline is FIFO. Packets arrived in the same time slot (from different flows) are queued randomly. The service time of each packet is independent and identically distributed, denoted as S . The following theorem provides bounds on the total timing capacity of such a system for service time distributions with a finite second moment.

PROPOSITION 5. The timing capacity of a server shared by N flows satisfies¹

$$(2.8) \quad \frac{B-1}{B}\mu \left(B - \log \left(1 + \frac{(B-1)\mu^2 E(S^2)}{2} \right) \right) \leq C \leq \mu B + 1,$$

where $\mu = 1/E(S)$. In particular, if the service-time distribution is geometrically distributed, then

$$\mu(B-1) - \mu \log B \leq C \leq \mu B + 1.$$

PROOF. As in the continuous-time case, we obtain a lower bound by computing the effective service time. However, it is difficult to precisely characterize the effective service time for a general server queue. To overcome the difficulty, we use the “waiting time + service time” as an upper bound on the effective service time. The intuition is that when the arrival rate is small, the inter-arrival of packets from the same flow is large, and thus the “waiting time + service time” becomes a good upper bound on the effective service time.

We first calculate an upper bound on the average effective service time. In [BK93], the average waiting time, W , is shown to be:

$$W = (W_r + W_s)/(1 - \rho),$$

where W_r is the average remaining service time, and W_s is the average service time of packets that arrive at the same slot, but served before the tagged packet. It is shown that

$$\begin{aligned} W_r &= \frac{E(S^2) - E(S)}{2E(S)}\rho = \frac{E(S^2)}{2E(S)}\rho - \frac{\rho}{2} \\ W_s &= \frac{E(K^2) - E(K)}{2E(K)}E(S), \end{aligned}$$

where S is the service time and K is the total number of packets arrived in a time-slot.

To obtain a lower bound, let the arrivals of each flow be Bernoulli with rate λ_0 . The aggregate arrival process is binomial with mean rate $N\lambda_0$, where N is the number of flows. Let $N\lambda_0 = (B-1)\mu/B$. We have

$$W_s = \frac{E(K^2) - E(K)}{2E(K)}E(S) = \frac{(N-1)\lambda_0}{2\mu} \leq \frac{\rho}{2}.$$

¹For the convenience of argument, the base of the logarithm in this section when we discuss discrete case is 2, and thus the unit of capacity is bits per slot.

Thus,

$$W = \frac{W_r + W_s}{1 - \rho} \leq \frac{\rho \frac{E(S^2)}{2E(S)}}{1 - \rho} = \frac{(B-1)E(S^2)}{2E(S)}.$$

Let $1/\mu'$ be the average effective service time. It satisfies

$$\frac{1}{\mu'} \leq \frac{1}{\mu} + W.$$

We have

$$\mu' \geq \frac{1}{\frac{1}{\mu} + \frac{(B-1)E(S^2)}{2E(S)}} = \frac{\mu}{1 + (B-1)\mu^2 E(S^2)/2}.$$

In [BA98], it is shown that the timing capacity of a geometric server has the smallest timing capacity among all service distributions with the same average service time, which is

$$C(\lambda) = h(\lambda) - \frac{\lambda}{\mu} h(\mu),$$

where $h(x) = -x \log x - (1-x) \log(1-x)$. To achieve this timing capacity, the arrival process is Bernoulli process with rate λ .

Thus, similar to the argument in the continuous case, we have

$$\begin{aligned} C &\geq N \left(h(\lambda_0) - \frac{\lambda_0}{\mu'} h(\mu') \right) \\ &= -N\lambda_0 \log \lambda_0 - N(1-\lambda_0) \log(1-\lambda_0) + N\lambda_0 \log \mu' + N\lambda_0 \frac{1-\mu'}{\mu'} \log(1-\mu') \\ &= N\lambda_0 \log \frac{\mu'}{\lambda_0} + N\lambda_0 \left[\frac{1-\lambda_0}{\lambda_0} \log \frac{1}{1-\lambda_0} - \frac{1-\mu'}{\mu'} \log \frac{1}{1-\mu'} \right]. \end{aligned}$$

We note that $f(x) = (1/x - 1) \log \frac{1}{1-x}$ is a decreasing function of x , $0 \leq x \leq 1$. Thus, when $\lambda_0 \leq \mu'$,

$$\begin{aligned} C(N) &\geq N\lambda_0 \log \frac{\mu'}{\lambda_0} \\ &\geq \frac{B-1}{B} \mu \log \left(\frac{\mu}{(1 + (B-1)\mu^2 E(S^2)/2)\lambda_0} \right) \\ &= \frac{B-1}{B} \mu \left(\log \frac{\mu}{\lambda_0} - \log(1 + (B-1)\mu^2 E(S^2)/2) \right) \\ &= \frac{B-1}{B} \mu \left(B + \log \frac{B}{B-1} - \log(1 + (B-1)\mu^2 E(S^2)/2) \right) \\ &= \frac{B-1}{B} \mu (B - \log(1 + (B-1)\mu^2 E(S^2)/2)). \end{aligned}$$

If $\lambda_0 \geq \mu'$, $N\lambda_0 \log \mu'/\lambda_0 \leq 0$, and thus the lower bound also holds.

The upper bound is clear:

$$I(X^n, A^n; X^n, D^n) \leq H(X^n, A^n) \leq H(X^n) + H(A^n) \leq \lambda B + h(\lambda) \leq \mu B + 1,$$

which completes the proof for the general service time distribution.

If the service time is geometrically distributed with mean $1/\mu$, then

$$E(S^2) = \frac{2(1-\mu)}{\mu^2} + \frac{1}{\mu} \leq \frac{2}{\mu^2}.$$

Substitute it into (2.8), and we have

$$\mu(B - 1) - \mu \log B \leq C \leq \mu B + 1,$$

which completes the proof. \square

In this section, we obtained asymptotically tight upper and lower bounds for the timing capacity of discrete queues with a general service time distribution. Note that we can also obtain a similar lower bound for continuous queues with general service time distributions as well. However, in continuous time queues, we have not been able to obtain an asymptotically tight upper bound.

3. Conclusions

In this paper, we first studied the case when a flow shares an exponential server with uncontrolled, independent interference traffic. When the interference traffic is Poisson, we obtained a lower bound on the timing capacity. Further, we derived asymptotically tight lower and upper bounds on the total timing capacity of multiple flows when the service time distribution is exponential. Lower and upper bounds are also calculated for discrete-time queues with a general service time distribution. We also studied the system where an eavesdropper can observe the sequence and contents of the departure packets, but not their timing. We showed that the amount of uncertainty at the eavesdropper is affected by the scheduling discipline at the server.

References

- [AV96] V. Anantharam and S. Verdú, *Bits through queues*, IEEE Transactions on Information Theory **42** (1996), no. 1, 4–18.
- [BA98] A. Bedekar and M. Azizoglu, *The information-theoretic capacity of discrete-time queues*, IEEE Transactions on Information Theory **44** (1998), no. 2, 446–461.
- [BG87] D. Bertsekas and R. Gallager, *Data networks*, Prentice-Hall Inc., 1987.
- [BK93] H. Bruneel and B. Kim, *Discrete-time models for communication systems including ATM*, Kluwer Academic Publishers, 1993.
- [CK78] I. Csiszar and J. Korner, *Broadcast channels with confidential messages*, IEEE Transactions on Information Theory **24** (1978), no. 3, 339–348.
- [Kle75] L. Kleinrock, *Queueing systems, volume i: Theory*, vol. I, Wiley Interscience, New York, 1975.
- [Sun99] R. Sundaresan, *Coded communication over timing channels*, Ph.D. thesis, Princeton University, 1999.
- [VH94] S. Verdú and T. Han, *A general formula for channel capacity*, IEEE Transactions on Information Theory **40** (1994), no. 4, 1147–1157.
- [Wyn75] A. Wyner, *The wire-tap channel*, Bell System Technical Journal **54** (1975), no. 8, 1355–1387.

COORDINATED SCIENCE LABORATORY, UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN, URBANA, IL 61801

E-mail address: xinliu@uiuc.edu

COORDINATED SCIENCE LABORATORY, UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN, URBANA, IL 61801

E-mail address: rsrikant@uiuc.edu