An Analytical Model Based on G/M/1 with Self-Similar Input to Provide End-to-End QoS in 3G Networks

Mohsin Iftikhar
School of Information Technologies
The University of Sydney
NSW, Sydney, 2006, Australia
+61 400 414 972
mohsinif@it.usyd.edu.au

Bjorn Landfeldt
School of Information Technologies
The University of Sydney
NSW, Sydney, 2006, Australia
+61 2 9351 8962
bjornl@it.usyd.edu.au

Mine Caglar
Department of Mathematics
Koc University
Istanbul, Turkey
+90 212 3381315
mcaglar@ku.edu.tr

ABSTRACT

The dramatic increase in demand for wireless Internet access has lead to the introduction of new wireless architectures and systems including 3G, Wi-Fi and WiMAX. 3G systems such as UMTS and CDMA2000 are leaning towards an all-IP architecture for transporting IP multimedia services, mainly due to its scalability and promising capability of inter-working heterogeneous wireless access networks. During the last ten years, substantial work has been done to understand the nature of wired IP traffic and it has been proven that IP traffic exhibits self-similar properties and burstiness over a large range of time scales. Recently, because of the large deployment of new wireless architectures, researchers have focused their attention towards understanding the nature of traffic carried by different wireless architecture and early studies have shown that wireless data traffic also exhibits strong long-range dependency. Thus, the classical tele-traffic theory based on a simple Markovian process cannot be used to evaluate the performance of wireless networks. Unfortunately, the area of understanding and modeling of different kinds of wireless traffic is still immature which constitutes a problem since it is crucial to guarantee tight bound QoS parameters to heterogeneous end users of the mobile Internet. In this paper, we make several contributions to the accurate modeling of wireless IP traffic by presenting a novel analytical model that takes into account four different classes of self-similar traffic. The model consists of four queues and is based on a G/M/1 queueing system. We analyze it on the basis of priority with no preemption and find exact packet delays. To date, no closed form expressions have been presented for G/M/1 with priority.

Categories and Subject Descriptors

C.2.3. [**Computer Communication Networks**]: Network Management G.3. [**Mathematics of Computing**]: Queueing Theory, Markov Processes

General Terms

Theory, Design, Performance, Reliability

Keywords

QoS, 3G, UMTS, GGSN, Self-Similar

1. INTRODUCTION

During the past decade, researchers have made significant efforts to understand the nature of Internet traffic and it has been proven that Internet traffic exhibits self-similar properties. The first study, which stimulated research on self-similar traffic, was based on measurements of Ethernet traffic at Bellcore [1]. Subsequently, the self-similar feature has been discovered in many other types of Internet traffic including studies on Transmission Control Protocol (TCP) [2, 3], WWW traffic [4], VBR video [5] and Signaling System No 7 [6]. Deeper studies into the characteristics of Internet traffic has discovered and investigated various properties such as self-similarity [7], long-range dependence [8] and scaling behavior at small time-scale [9]. The references [10, 11] provide two extensive bibliographies on self-similarity and long-range dependence research covering both theoretical and applied papers on the subject.

Concurrently, over the past few years, we have witnessed a growing popularity of Third Generation Systems (3G), which have been designed to provide high-speed data services and multimedia applications over mobile personal communication networks. The Universal Mobile Telecommunication System (UMTS) is the predominant global standard for 3G developed by Third Generation Partnership Project (3GPP) [12]. The UMTS architecture is shown in Fig. 1. It consists of two service domains, a Circuit Switched (CS) service domain and a Packet Switched (PS) service domain, which is of interest in this paper. In the PS service domain, a UMTS network connects to a public data network (PDN) through Serving GPRS Support node (SGSN) and Gateway GPRS support node (GGSN). 3GPP has defined four different QoS classes for UMTS; (1) Conversational (2) Interactive (3) Streaming and (4) Background, conversational being the most delay-sensitive and background the least delay sensitive class [12].

With the increasing demand of Internet connectivity and the flexibility and wide deployment of IP technologies, there has emerged a paradigm shift towards IP-based solutions for wireless networking [13]. Several Wireless IP architectures have been proposed [17-23] based on three main IP QoS models, IntServ [14], DiffServ [15] and MPLS [16]. 3GPP has also recently introduced a new domain called IP Multimedia Subsystem (IMS) for UMTS. The main objective of IMS is to deliver innovative

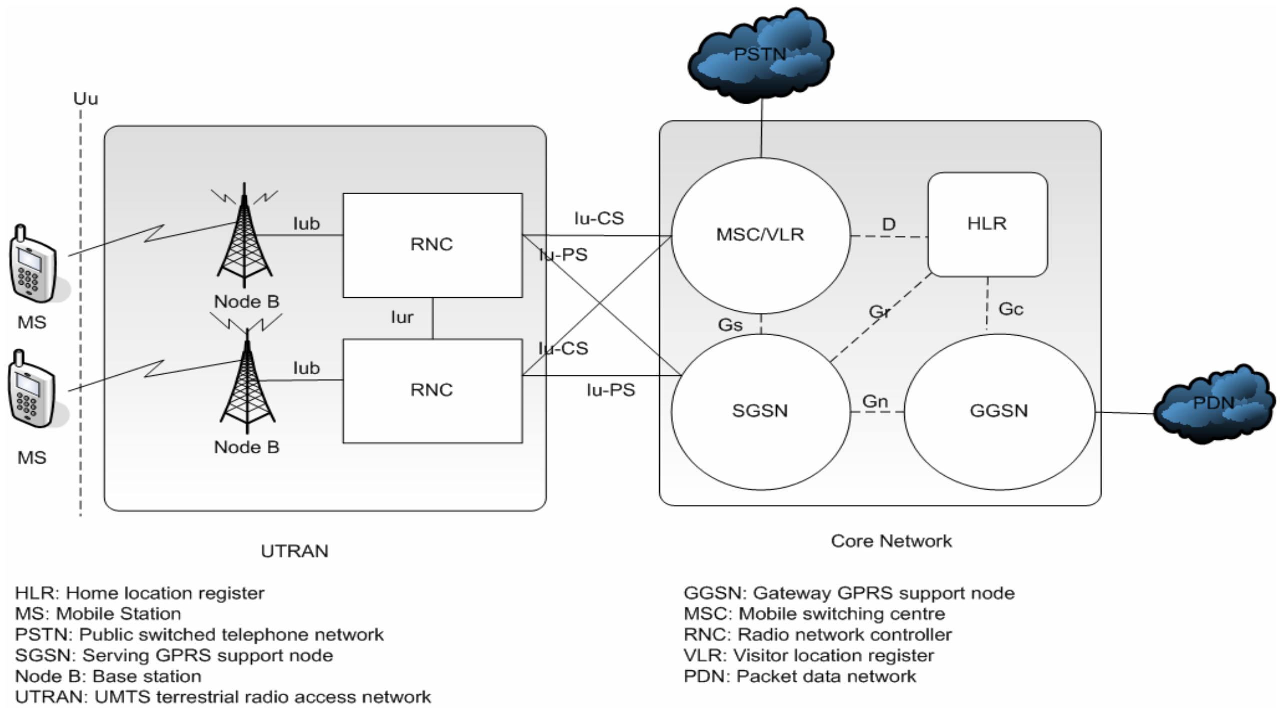

Fig. 1: A Simplified UMTS Network Architecture

and cost-effective services such as IP telephony, media streaming and multiparty gaming by providing IP connectivity to every mobile device [24].

In the light of this, researchers have recently focused on understanding the nature of wireless IP traffic and early studies have shown that wireless data traffic also exhibits self-similarity and long-range dependency [25-28]. Much of the current understanding of wireless IP traffic modeling is based on the simplistic Poisson model, which can yield misleading results and hence poor wireless network planning. Since the properties and behavior of self-similar traffic is very different from traditional Poisson or Markovian traffic, several issues need to be addressed in modeling wireless IP traffic to provide end-to-end QoS to a variety of heterogeneous applications. We begin by giving an overview of related work on wired and wireless IP traffic modeling along with a comparison of our model with previous work.

2. RELATED WORK

In this section, we first discuss the related work which has been done in the area of performance evaluation of wired IP and Wireless IP networks under self-similar input and then we compare our model with the previous ones.

2.1 Previous Work on IP Traffic Modeling

There has been much work done on Internet traffic modeling based on queueing theory in the presence of self-similar traffic [29-34]. In [33], a Matrix Geometric (analytical) method is used to compute numerical results for a two class DiffServ link being fed by a Markovian Modulated Poisson Process (MMPP) input. A weakness of this model is that MMPP may require an estimation of a large number of parameters. An OPNET based simulation approach was adopted in [34] to see the impact of self-similarity on the performance evaluation of DiffServ networks. As a result, an idea of expected queue length was given in relation to the Hurst parameter and server utilization. It is difficult to offer guaranteed QoS parameters on the basis of such analysis. The major weakness of the majority of available queueing based results is that only the FIFO queueing discipline has been considered for serving the incoming traffic and thus differential treatment to different kinds of traffic can not be provided. In addition, the previous results are asymptotic. We also refer the readers to [35-39] for an overview of previous work that has been carried out to evaluate the performance of IP networks. The major drawback of the existing work is that, the queueing models considered are not able to capture the self-similar characteristics of Internet traffic. Furthermore, it is important to note that most of the previous work is focused on the analysis of one type of traffic only without discussing its affect on the performance of other kinds of network traffic.

2.2 Previous Work on Wireless IP Traffic Modeling

Few studies have focused on wireless traffic modeling and here we discuss the most relevant work. As shown in Fig. 1, the principle of allocation of data flows between end users and GGSN leads to increasing load on the network elements when moving closer to the GGSN. Hence, GGSN is the node most exposed to self-similar influence in UMTS [40]. The influence of self-similar input on GGSN performance in the UMTS Release 5 IM-subsystem has been analyzed on the basis of a FBM/D/1/W queueing system (FBM-Fractional Brownian Motion) in [40]. In this work, different probabilistic parameters of GGSN such as average queue length and average service rate were also found. The work in [41] presents modeling and a simulation study of the Telus Mobility (a commercial service provider) Cellular Digital Packet Data (CDPD) network. The collected results on average queueing delay and buffer overflow probability indicated that genuine traffic traces produce longer queues as compared to traditional Poisson based traffic models. To get an overview of the analysis done in wireless IP traffic modeling with self-similar input, we refer the readers to [42-45]. These studies are merely based on characterization of wireless traffic. To provide differential treatment to multiple traffic classes with different QoS demands, there is a need to accurately determine end-to-end QoS parameters such as delay, jitter, throughput, packet loss, availability and per-flow sequence preservation.

2.3 Comparison of our Model with Prior Work

To overcome the limitations of the previous work in traffic modeling (wired and wireless IP traffic), we present a realistic and novel analytical model by considering four different classes of traffic that exhibit long-range dependence and self-similarity. Our model implements four queues based on a G/M/1 queueing system and we analyze it on the basis of priority with no preemption. The traffic model considered is parsimonious with few parameters and has been studied in [46]. The model is furthermore similar to on/off processes, in particular to its variation N-Burst model studied in [47] where packets are incorporated. However, only a single type of traffic is considered in [47]. We present a novel analytical approach and make the following contributions to Wireless IP traffic modeling.

Interarrival Time Calculations: We calculate the packet interarrival time distributions for the particular self-similar traffic model [46] for the first time in this paper. The distribution of cross interarrival time between different types of packets is derived on the basis of single packet results.

Packet Delays for Multiple Self-Similar Traffic Classes: We consider a G/M/1 queueing system which takes into account four different classes of self-similar input traffic denoted by SS/M/1 and analyze it on the basis of non preemptive priority and find exact packet delays. To date, no closed form expressions have been presented for G/M/1 with priority.

Embedded Markov Chain Formulation: We also formulate the embedded Markov chain of G/M/1 by considering all possible states and derive the corresponding transition probabilities.

The rest of the paper is organized as follows. Section 3 and 4 are devoted to explaining the self-similar traffic model with multiple classes and the calculation of interarrival times respectively. Section 5 explains the procedure of formulating the embedded Markov Chain along with the derivation of packet delays. The applications of the model are discussed in section 6. Finally, conclusion and future work is given in Section 7.

3. TRAFFIC MODEL

The traffic model considered here [46] belongs to a particular class of self-similar traffic models also called telecom process in [48], recently. The model captures the dynamics of packet generation while accounting for the scaling properties of the traffic in telecommunication networks. Such models, also called infinite source models, are similar to on/off processes with heavy tailed on and/or off times. What is more, our model abstracts the packet arrival process in particular and facilitates queueing analysis by the approaches developed in the sequel.

In the framework of a Poisson point process, the model represents an infinite number of potential sources. The traffic is found by aggregating the number of packets generated by such sources. Each source initiates a session with a heavy-tailed distribution, in particular a Pareto distribution whose density is given by

$$g(r) = \delta b^{\delta} r^{-\delta-1}, \qquad r > b$$

where δ is related to the Hurst parameter by $H = (3-\delta)/2$. The sessions arrive according to a Poisson process with rate λ. The packets arrive according to a Poisson process with rate α, locally, over each session.

For each class, the traffic $Y(t)$ measured as the total number of packets injected in [0, t] is found by

$$Y(t) = \sum_{S_i \le t} U_i (R_i \wedge (t - S_i))$$

where U_i, R_i, S_i denote the local Poisson process, the duration and the arrival time of session i, respectively. Hence, $Y(t)$ corresponds to the sum of packets generated by all sessions initiated in [0,t] until the session expires if that happens before t, and until t if is does not. The stationary version of this model based on an infinite past is considered in calculations below. The packet sizes are assumed to be fixed because each queue corresponds to a certain type of application where the packets have fixed size or at least fixed service time distribution.

The traffic model Y is long-range dependent and almost second-order self-similar; the auto covariance function of its increments is that of fractional Gaussian noise. Three different heavy traffic limits are possible depending on the rate of increase in the traffic parameters [46, 48]. Two of these limits are well known self-similar processes, fractional Brownian motion and Levy process, which do not account for packet dynamics in particular.

4. INTERARRIVAL TIMES

Packet interarrival time distributions for the particular self-similar traffic model are calculated for the first time in this paper. We consider a single type of packet first. The distributions of cross interarrival time between different types of packets are derived on the basis of single packet results.

4.1 Interarrival Times for a Single Class

Although the packet arrival process itself is long-range dependent and shows self-similarity, the number of alive sessions at a period of time, say of length t, has a stationary distribution and is Poisson distributed. The alive sessions at any time can be further split into independent components as those session that last longer than t and those that expire before t. Such results are well known [49, pg.273] and will be used to derive the interarrival time distribution of the packets.

Given that there is a packet arrival at an instant in time, we aim to find the distribution of the time until next arrival denoted by T. We will find $\overline{F}(t) = P\{T > t\}$, for $t \ge 0$. When the event $\{T > t\}$ is considered, the information that there is a packet arrival is equivalent to the information that there is at least one session alive at the given instant. This follows from the assumption that local packet generation process is Poisson over each session. The probability that next interarrival is greater than t

on a particular session is the same as the probability that the remaining time until next arrival is greater than t due to the memoryless property of exponential distribution. That is,

$P\{T > t\} =$ P {Next packet interarrival is greater than t | there is a packet arrival}

$= \frac{1}{\rho}$ P {Next packet interarrival is greater than t, there is at least one alive session}

where ρ is the probability that there is at least one alive session, in other words the utilization of an $M/G/\infty$ system. The event that next packet interarrival is greater than t can be split as follows:

- The active sessions that expire before t do not incur any new arrivals.
- The active sessions that expire after t do not incur any new arrivals
- No new session arrivals in t or at least one session arrival with no packet arrival in t.

We find the probability that all three events occur at the same time by using the independence of a Poisson point process over disjoint sets. The result is

$$P\{T > t\} = \frac{1}{\rho}\{e^{-\nu(A_t)}e^{-\nu(B_t)}\exp[-\lambda t(1-e^{-\alpha t})]$$
$$(\exp[\nu(A_t)e^{-\alpha t}]-1)\exp[\nu(B_t)(1-e^{-\alpha t})/\alpha t]-1)$$
$$+\mathrm{e}^{-\nu(A_t)}\mathrm{e}^{-\nu(B_t)}\exp[-\lambda t(1-e^{-\alpha t})](\exp[\nu(A_t)\mathrm{e}^{-\alpha t}]-1)$$
$$+\mathrm{e}^{-\nu(A_t)}\mathrm{e}^{-\nu(B_t)}\exp[-\lambda t(1-e^{-\alpha t})](\exp[\nu(B_t)(1-\mathrm{e}^{-\alpha t})/\alpha t]-1)\}$$

where

$$\nu(A_t) = \lambda\int_t^{\infty}(y-t)g(y)\,dy \qquad (1)$$

$$\nu(B_t) = \lambda\left[\int_0^t yg(y)\,dy + t\,\overline{G}(t)\right] \qquad (2)$$

and

$$\rho = 1-\exp(-\lambda\,\mathrm{E}[\text{session duration}])$$
$$= 1-\exp[-\lambda\,\delta b/(\delta-1)]$$

because the steady state number in the system in $M/G/\infty$ queue is Poisson distributed with mean λ E[Session duration] [50], and δ and b are the parameters of the session duration with complementary distribution function $\overline{G}$ and density

$$g(r) = \delta b^{\delta} r^{-\delta-1} \qquad r > b$$

which is Pareto.

4.2 Interarrival Times for Multiple Classes

Here we explain the detailed procedure to find out the Interarrival times for two classes, the Interarrival times for more than two classes can be found in a similar way. Let T_{ij} denote the interarrival time between a class i packet that comes first and a class j packet that follows, $i, j = 1,2.$ The analysis, which can be extended to $i, j \geq 3$, provides a method for other self-similar models as well provided that the distribution of interarrivals T_i are available.

For the consecutive packet 1 arrival time T_{11}, we have

$$P\{T_{11} > t\} = P\{T_1 > t, \text{ no arrivals of class 2 in } T_1\}$$

$$= \int_t^{\infty} P\{\text{no arrivals of class 2 in s}\}\, f_{T_1}(s)\,ds$$

$$= \int_t^{\infty} \overline{F}_2^{\,0}(s)\, f_{T_1}(s)\,ds$$

where $\overline{F_2}^{\,0}(t) = e^{-\nu_2(B_t)}e^{-\nu_2(A_t)}$.

$$\exp[-\lambda_2 t(1-e^{-\alpha_2 t})]\exp[\nu_2(A_t)e^{-\alpha_2 t}]$$
$$\exp[\nu_2(B_t)(1-e^{-\alpha_2 t})/\alpha_2 t] \qquad (3)$$

$\nu_2(A_t)$ and $\nu_2(B_t)$ are defined analogously as in (1) and (2), and we used the independence of class 1 and 2 packet inputs. Here, $\overline{F_2}^{\,0}$ is found through similar arguments used for P {T>t} in the last subsection, without assuming that there is an alive session of type 2 . As a result, by differentiation we find

$$f_{T_{11}}(t) = f_{T_1}(t)\,\overline{F}_2^{\,0}(t)$$

Now consider the interarrival time T_{12} occurring between a class 1 packet followed by a class 2 packet. For T_{12}, we get

$$P\{T_{12} \leq t\} =$$

$$\int_0^t f_2^0(s) P\{\text{no arrivals of class1 in } s \mid \text{a class1 packet arrived}\}ds$$

$$= \int_0^t f_2^0(s)\overline{F}_{T_1}(s)ds$$

where $f_2^0(s)$ is the density function corresponding to the event that there is an arrival of class 2 packet at time s, and we used independence of class 1 and 2 packet streams. As a matter of fact, $f_2^0(s)$ can be obtained by taking the derivative of the complementary distribution function $\overline{F}_2^{\,0}$ given in (3). As a result, we get $f_{T_{12}}(t) = f_2^0(t)\,\overline{F}_{T_1}(t)$

Similarly, it can be shown that

$$f_{T_{22}}(t) = f_{T_2}(t)\,\overline{F}_1^{\,0}(t)\,,\;\; f_{T_{21}}(t) = f_1^{\,0}(t)\,\overline{F}_{T_2}(t)$$

5. QUEUEING MODEL

We consider a model of four queues based on G/M/1 by considering four different classes of self-similar input traffic denoted by SS/M/1, and analyze it on the basis of priority with no preemption. Let the service time distribution have rate μ_1, μ_2, μ_3 and μ_4 for type 1, type 2, type 3 and type 4 packets, respectively, and let type 1 packets have the highest priority and type 4 packets have the lowest priority.

5.1 $SS/M/1$ With Four Classes

The usual embedded Markov chain [51] formulation of $G/M/1$ is based on the observation of the queueing system at the time of arrival instants, right before an arrival. At such instants, the number in the system is the number of packets that arriving packet sees in the queue plus packets in service, if any, excluding the arriving packet itself. We specify the states and the transition probability matrix *P* of the Markov chain with the self-similar model for four types of traffic.

Let $\{X_n : n \geq 0\}$ denote the embedded Markov chain at the time of arrival instants. As the service is based on priority, the type of packet in service is important at each arrival instant of a given type of packet to determine the queueing time. Therefore, we define the state space as:

$$\begin{aligned} &S = \{(i_1, i_2, i_3, i_4, a, s) : a \in \{a_1, a_2, a_3, a_4\}, \\ &s \in \{s_1, s_2, s_3, s_4, I\}, i_1, i_2, i_3, i_4 \in Z_+\} \end{aligned} \qquad (4)$$

where a_1, a_2, a_3, a_4 are labels to denote the type of arrival, s_1, s_2, s_3, s_4 are labels to denote the type of packet in service, i_1, i_2, i_3, i_4 are the number of packets in each queue including a possible packet in service, I denotes the idle state in which no packet is in service or queued and Z_+ is the set of nonnegative integers. Some of the states in the state space *S* given in (4) have zero probability. For example, $(i_1, 0, i_3, i_4, a_1, s_2)$ is impossible. The particular notation in (4) for *S* is chosen for simplicity, although the impossible states could be excluded from *S*. Each possible state, the reachable states from each and the corresponding transition probabilities will be calculated.

5.2 States of the Embedded Markov Chain

The states of the Markov chain and the possible transitions with respective probabilities can be enumerated by considering each case. We will only analyze the states with non-empty queues in this paper.

5.2.1 States $(i_1, i_2, i_3, i_4, a, s)$ *with* $i_1, i_2, i_3, i_4 \neq 0$

We can divide the states and transitions into 256 groups. Because (*a, s*) can occur 4x4=16 different ways, and the next state (*p, q*) can be composed similarly in 16 different ways as $a, p \in \{a_1, a_2, a_3, a_4\}$ and $s, q \in \{s_1, s_2, s_3, s_4\}$. We will analyze only the first one in detail; the others follow similarly.

5.2.2 Transition from $(i_1, i_2, i_3, i_4, a_1, s_1) \rightarrow (j_1, j_2, j_3, j_4, a_2, s_2)$

This is the case where a transition occurs from an arrival of type 1 to an arrival of type 2 such that the first arrival has seen a type 1 packet in service, i_1 packets of type 1 (equivalently, total of queue 1 and the packet in service) and i_2 packets of type 2 (in this case only queue 2), i_3 packets of type 3 and i_4 packets of type 4 in the system. The transition occurs to j_1 packets of type 1, j_2 packets of type 2, with a type 2 packet in service, j_3 packets of type 3 and j_4 packets of type 4 in the system. Due to priority scheduling, an arrival of type 2 can see a type 2 packet in service in the next state only if all type 1 packets including the one that arrived in the previous state are exhausted during the interarrival time. That is why j_1 can take only the value 0 and exactly $i_1 + 1$ packets of type 1 are served. In contrast, the number of packets served from queue 2, say *k*, can be anywhere between 0 and $i_2 - 1$ as at least one type 2 packet is in the system, one being in service, when a new arrival occurs. The transition probabilities are

$$P\{X_{n+1} = (0, i_2 - k, i_3, i_4, a_2, s_2) \mid X_n = (i_1, i_2, i_3, i_4, a_1, s_1)\}$$

$= P\{i_1 + 1$ served from type 1, *k* served from type 2 and a type 2 packet remains in service during T_{12} }

where we use the fact that the remaining service time of a type 1 packet in service has the same exponential distribution Exp(μ_1), due to the memory-less property of a Markovian service. Therefore, for $k = 0, \ldots, i_2 - 1$

$$P\{X_{n+1} = (0, i_2 - k, i_3, i_4, a_2, s_2) \mid X_n = (i_1, i_2, i_3, i_4, a_1, s_1)\}$$

$$= \int_0^{\infty}\int_0^{t}\int_{t-x}^{\infty} f_{S_2}(s)\, f_{S_1^{i_1+1} + S_2^{k}}(x)\, f_{T_{12}}(t)\, ds\,dx\,dt$$

where S_m^l : sum of *l* independent service times of type *m* packets, *m*=1, 2, $l \in Z_+$. Note that S_m^l has an Erlang distribution with parameters (l, μ_m) as each service time has an exponential distribution, and the sum $S_1^{l_1} + S_2^{l_2}$ being the sum of several exponentially distributed random variables has a hypoexponential distribution. The density functions of all these distributions can easily be evaluated numerically. Similarly, we can enumerate all 256 cases. The results for first 64 cases are given in Table 1 in the Appendix.

5.3 Limiting Distribution and Waiting Times

Steady state distribution π as seen by an arrival can be found by solving $\pi P = \pi$ using the transition matrix P of the Markov chain analyzed above. In practice, the queue capacity is limited in a router. So, the steady state distribution exists.

To the best of our knowledge, no previous analytical expressions are available for the waiting time of a G/M/1 queue with priority. Our analysis relies on the limiting distribution of the state of the queue at the arrival instances, which can be computed using the analysis given above for our self-similar traffic model. In general, the following analysis is valid for any G/M/1 queueing system where the limiting distribution π at the arrival instances can be computed. The expected waiting time for the highest priority queue can be found as

$$E[W_1] = \sum_{j_1=1}^{J_1-1}\sum_{j_2=0}^{J_2}\sum_{j_3=0}^{J_3}\sum_{j_4=0}^{J_4}\frac{j_1}{\mu_1}\pi(j_1,j_2,j_3,j_4,a_1,s_1)+$$
$$\sum_{j_1=0}^{J_1-1}\sum_{j_2=1}^{J_2}\sum_{j_3=0}^{J_3}\sum_{j_4=0}^{J_4}\left(\frac{j_1}{\mu_1}+\frac{1}{\mu_2}\right)\pi(j_1,j_2,j_3,j_4,a_1,s_2)+$$
$$\sum_{j_1=0}^{J_1-1}\sum_{j_2=0}^{J_2}\sum_{j_3=1}^{J_3}\sum_{j_4=0}^{J_4}\left(\frac{j_1}{\mu_1}+\frac{1}{\mu_3}\right)\pi(j_1,j_2,j_3,j_4,a_1,s_3)+$$
$$\sum_{j_1=0}^{J_1-1}\sum_{j_2=0}^{J_2}\sum_{j_3=0}^{J_3}\sum_{j_4=1}^{J_4}\left(\frac{j_1}{\mu_1}+\frac{1}{\mu_4}\right)\pi(j_1,j_2,j_3,j_4,a_1,s_4)$$

where J_1, J_2, J_3 and J_4 are the respective capacities of each queue. This follows clearly from the fact that an arriving packet of higher priority will wait until all packets of the same priority as well as the packet in service are served. Depending on the type of the packet in service, we have the constituent expressions in the sum.

On the other hand, we obtain the expected waiting time for the low priority queues by analyzing the events that constitute this delay. The amount of work in the system at any time is defined as the (random) sum of all service times that will be required by the packets in the system at that instant. The waiting time of a type 2 packet (which is 2nd highest priority queue) can be written as

$$W_2 = Z_1 + Z_2 + Z_3 + \dots \qquad (5)$$

where Z_1 is the amount of work seen by the arriving packet in queue 1 and queue 2 (i.e, higher priority and equal priority), Z_2 is the amount of work associated with higher priority (i.e.type 1) packets arriving during Z_1, Z_3 is the amount of work associated with type 1 packets arriving during Z_2, and so on. As illustrated in Fig.2, the waiting time of an arriving packet of type 2 is indeed given by the total workload building in front of it. The arrows in the figure denote the arrival times of type 1 packets, and all the oblique lines have 45 degrees angle with the time axis. In this figure the waiting time is

$W_2 = Z_1 + Z_2 + Z_3 + Z_4$ for example.

Let M_j denote the number of type j arrivals over Z_i, j=1,2,….Then

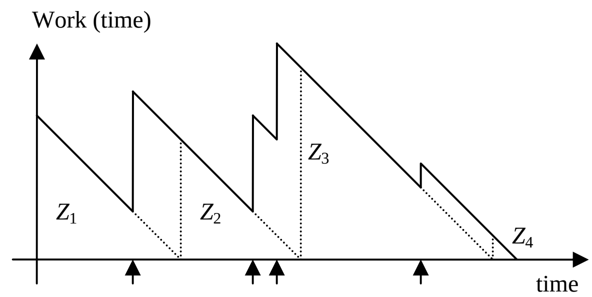

Fig.2 Waiting time of a type 2 packet in terms of Z_j's.

$$W_2 = Z_1 + S_1^{M_1} + S_1^{M_2} + \cdots$$

where $S_1^{M_j}$ denotes the random sum of M_j independent service times of type 1 packets. Then,

$$E[W_2] = E[Z_1] + E[S_1]E[M_1] + E[S_1]E[M_2] + \cdots$$

since the service times and the arrival process are independent. For a stationary packet arrival process, we get

$$E[M_j] = E[E[M_j \mid Z_j]] = E[c_1 Z_j] = c_1 E[Z_j]$$

due to mentioned independence, where $c_1 > 0$ is a constant particular to the arrival process. That is, expectation of the number of arrivals in any period of time is proportional to the length of that period because of stationarity in time and linearity of expectation. In our stationary self-similar traffic input process, c_1 is the expected number of arrivals per unit time which can be called the arrival rate, given by the product of the arrival rate of session arrivals, the arrival rate of packets over a session, and the expected session length [46]. Explicitly, $c_1 = \lambda\alpha\delta b/(\delta-1)$. Hence, the expected waiting time reduces to

$$E[W_2] = E[Z_1] + E[S_1]c_1E[Z_1] + E[S_1]c_1E[Z_2] + \cdots$$
$$= E[Z_1] + \frac{c_1}{\mu_1}(E[Z_1] + E[Z_2] + \cdots)$$
$$= E[Z_1] + \frac{c_1}{\mu_1}E[W_2]$$

In view of (5), therefore we get $E[W_2]$ from

$$E[W_2] = \sum_{j_1=1}^{J_1}\sum_{j_2=0}^{J_2-1}\sum_{j_3=0}^{J_3}\sum_{j_4=0}^{J_4}\left(\frac{j_1}{\mu_1}+\frac{j_2}{\mu_2}\right)\pi(j_1,j_2,j_3,j_4,a_2,s_1)+$$
$$\sum_{j_1=0}^{J_1}\sum_{j_2=1}^{J_2-1}\sum_{j_3=0}^{J_3}\sum_{j_4=0}^{J_4}\left(\frac{j_1}{\mu_1}+\frac{j_2}{\mu_2}\right)\pi(j_1,j_2,j_3,j_4,a_2,s_2)+$$
$$\sum_{j_1=0}^{J_1}\sum_{j_2=0}^{J_2-1}\sum_{j_3=1}^{J_3}\sum_{j_4=0}^{J_4}\left(\frac{j_1}{\mu_1}+\frac{j_2}{\mu_2}+\frac{1}{\mu_3}\right)\pi(j_1,j_2,j_3,j_4,a_2,s_3)+$$
$$\sum_{j_1=0}^{J_1}\sum_{j_2=0}^{J_2-1}\sum_{j_3=0}^{J_3}\sum_{j_4=1}^{J_4}\left(\frac{j_1}{\mu_1}+\frac{j_2}{\mu_2}+\frac{1}{\mu_4}\right)\pi(j_1,j_2,j_3,j_4,a_2,s_4)+\frac{c_1E[W_2]}{\mu_1}$$

Similarly, we can directly write down the expected waiting time for a packet of type 3 (3rd priority queue) and type 4 (lowest priority queue). The expected waiting time for a packet of type 3 can be found from:

$$E[W_3] = \sum_{j_1=1}^{J_1}\sum_{j_2=0}^{J_2}\sum_{j_3=0}^{J_3-1}\sum_{j_4=0}^{J_4}(\frac{j_1}{\mu_1}+\frac{j_2}{\mu_2}+\frac{j_3}{\mu_3})\pi(j_1, j_2, j_3, j_4, a_3, s_1) +$$

$$\sum_{j_1=0}^{J_1}\sum_{j_2=1}^{J_2}\sum_{j_3=0}^{J_3-1}\sum_{j_4=0}^{J_4}(\frac{j_1}{\mu_1}+\frac{j_2}{\mu_2}+\frac{j_3}{\mu_3})\pi(j_1, j_2, j_3, j_4, a_3, s_2) +$$

$$\sum_{j_1=0}^{J_1}\sum_{j_2=0}^{J_2}\sum_{j_3=1}^{J_3-1}\sum_{j_4=0}^{J_4}(\frac{j_1}{\mu_1}+\frac{j_2}{\mu_2}+\frac{j_3}{\mu_3})\pi(j_1, j_2, j_3, j_4, a_3, s_3) +$$

$$\sum_{j_1=0}^{J_1}\sum_{j_2=0}^{J_2}\sum_{j_3=0}^{J_3-1}\sum_{j_4=1}^{J_4}(\frac{j_1}{\mu_1}+\frac{j_2}{\mu_2}+\frac{j_3}{\mu_3}+\frac{1}{\mu_4})\pi(j_1, j_2, j_3, j_4, a_3, s_4) +$$

$$(\frac{c_1}{\mu_1}+\frac{c_2}{\mu_2})E[W_3]$$

and $E[W_4]$ can be determined from

$$E[W_4] = \sum_{j_1=1}^{J_1}\sum_{j_2=0}^{J_2}\sum_{j_3=0}^{J_3}\sum_{j_4=0}^{J_4-1}(\frac{j_1}{\mu_1}+\frac{j_2}{\mu_2}+\frac{j_3}{\mu_3}+\frac{j_4}{\mu_4})\pi(j_1, j_2, j_3, j_4, a_4, s_1) +$$

$$\sum_{j_1=0}^{J_1}\sum_{j_2=1}^{J_2}\sum_{j_3=0}^{J_3}\sum_{j_4=0}^{J_4-1}(\frac{j_1}{\mu_1}+\frac{j_2}{\mu_2}+\frac{j_3}{\mu_3}+\frac{j_4}{\mu_4})\pi(j_1, j_2, j_3, j_4, a_4, s_2) +$$

$$\sum_{j_1=0}^{J_1}\sum_{j_2=0}^{J_2}\sum_{j_3=1}^{J_3}\sum_{j_4=0}^{J_4-1}(\frac{j_1}{\mu_1}+\frac{j_2}{\mu_2}+\frac{j_3}{\mu_3}+\frac{j_4}{\mu_4})\pi(j_1, j_2, j_3, j_4, a_4, s_3) +$$

$$\sum_{j_1=0}^{J_1}\sum_{j_2=0}^{J_2}\sum_{j_3=0}^{J_3}\sum_{j_4=1}^{J_4-1}(\frac{j_1}{\mu_1}+\frac{j_2}{\mu_2}+\frac{j_3}{\mu_3}+\frac{j_4}{\mu_4})\pi(j_1, j_2, j_3, j_4, a_4, s_4) +$$

$$(\frac{c_1}{\mu_1}+\frac{c_2}{\mu_2}+\frac{c_3}{\mu_3})E[W_4]$$

6. APPLICATIONS OF THE MODEL

Here we give an overview of the prime application of the model. 3G systems such as UMTS and CDMA2000 are leaning towards an all-IP network architecture for transporting IP multimedia services [52]. An all-IP DiffServ platform is the currently most promising architecture to interwork the heterogeneous wireless access networks and the Internet to provide broadband access, seamless global roaming and QoS guarantees for various IP multimedia services [53]. To transport UMTS services through IP networks without loosing end-to-end QoS provisioning, a consistent and efficient QoS mapping between UMTS services and IP QoS classes is required. According to 3GPP, UMTS-to-IP QoS mapping is performed by a translation function in the GGSN router that classifies each UMTS packet flow and maps it to a suitable IP QoS class [52]. In order to make accurate mappings and to ensure guaranteed QoS parameters to the end user of mobile Internet, it is essential to being able to accurately model the end-to-end behavior of different classes of wireless IP traffic (conversational, interactive, streaming and background) passing through a DiffServ domain. Several queueing tools have been developed that can be implemented in IP routers within different QoS domains including Priority Queueing (PQ), Custom Queueing (CQ), Weighted Fair Queueing (WFQ), Class Based Weighted Fair Queueing (CBWFQ) and Low-Latency Queueing (LLQ) [54].

Our model is directly applicable to the problem of determining the end-to-end queueing behavior of IP traffic through both Wired and wireless IP domains, but modeling accuracy is more crucial in resource constrained environments such as wireless networks. For example, our model is directly able to analyze the behavior of four different QoS classes of UMTS traffic passing through a DiffServ domain, in which routers are implemented with priority queueing. Thus, the model enables tighter bounds on actual behavior so that over-provisioning can be minimized. It also enables translations of traffic behavior between different kinds of QoS domains so that it is possible to map reservations made in different domains to provide session continuity.

7. CONCLUSION AND FUTURE WORK

In this paper, we have presented a novel analytical model based on G/M/1 queueing system for accurate modeling of wireless IP traffic behavior under the assumption of four different classes of self-similar traffic. We have analyzed it on the basis of non-preemptive priority and explicit expressions of expected waiting time for the corresponding classes have been derived. The model represents an important step towards the overall aim of finding realistic (under self-similar traffic assumptions) end-to-end QoS behavior (in terms of QoS parameters such as delay, jitter and throughput) of multiple traffic classes passing through heterogeneous wireless IP domains (IntServ, DiffServ and MPLS). At the moment, we are working on the numerical analysis to find solutions to the suggested imbedded Markov Chain in order to find exact QoS parameter bounds for a given system. Our future work will focus to analyze the performance of different QoS domains implemented with different queueing disciplines.

8. REFERENCES

[1] W. Leland, M. Taqqu, W. Willinger and D. Wilson, "On the self-similar nature of Ethernet traffic (extended version)", *IEEE/ACM Transactions on Networking,* vol. 2. no. 1, pp. 1-15, Feb. 1994.

[2] V. Paxon, "Empirically derived analytical models of wide-area TCP connections", *IEEE/ACM Transactions on Networking*, vol. 2, pp. 316-336, Aug. 1994

[3] V. Paxon and S. Floyd , "Wide-area traffic: the failure of Poisson modeling", *in Proc. ACM SIGCOMM 94*, London, U.K., Aug. 1994, pp. 257-268

[4] M. Crovella and A. Bestavros, "Explaining World Wide Web Traffic Self-Similarity", *Tech. Rep. TR-95-015*, *Boston University, CS Dept, Boston*, MA 02215, Aug. 1995

[5] M. W. Garrett and W. Willinger, "Analysis, Modeling and Generation of Self-Similar VBR Video Traffic", *ACM Computer Communication Review*, vol. 24, Oct. 1994, SIGCOMM 94 Symposium

[6] W. Willinger et al, "Statistical analysis of CCSN/SS7 traffic data from working CCS subnetworks", *IEEE. Journal on Selected Areas of Communication*, vol. 12, no. 3, pp. 544-551, Apr. 1994

[7] M. Crovella and A. Bestavros, "Self-Similarity in World Wide Web Traffic: Evidence and Possible Causes", *in ACM Sigmetrics*, May 1996

[8] J.C Bolot and M. Grossglauser, "On the Relevance of Long-Range Dependence in Network Traffic", *Computer Communication Review*, vol. 26, no. 4, pp. 15-24, October 1996.

[9] Z. L. Zhang, V. Ribeiro, S. Moon and C. Diot, “Small-Time Scaling behavior of internet backbone traffic: An Empirical Study”, *in IEEE INFOCOM*, march 2003

[10] M. S. Taqqu, “Self-Similar processes”. In S. Kotz and N. Johnson, editors, *Encyclopedia of Statistical Sciences*, vol. 8, pp. 352-357. Wiley, New York, 1988.

[11] W. Willinger, M.S Taqqu and A. Erramilli, “A bibliographical guide to self-similar traffic and performance modeling for modern high speed networks”, In F. P. Kelly, S. Zachary and I. Ziedins, editors, *Stochastic Networks: Theory and Applications*, pp. 339-366, Claredon Press, Oxford, 1996

[12] H. Holma and A. Taskala, “WCDMA for UMTS, Radio Access for Third Generation Mobile Communications, 2nd Edition”, *John Wiley & Sons*, Ltd. 2002, pp. 1-5

[13] J. Yang and I. Kriaras, “Migration to all-IP based UMTS networks, “ *IEEE 1st International Conference on 3G Mobile Communication Technologies* , 27-29 March, 2000, pp. 19-23

[14] W. Stallings, “Integrated Services Architecture: The Next-Generation Internet”, *International Journal of Network Management*, 9, 1999, pp. 38-43

[15] S. Blake et al., “An Architecture for Differentiated Services”, *IETF RFC 2475*, Dec. 1998

[16] Rosen E. et al., “Multiprotocol Label Switching (MPLS) Architecture”, *RFC 3031*, Jan. 2001

[17] K. Venken, J. De Vriendt and D. De Vleeschauwer, “Designing a DiffServ-capable IP-backbone for the UTRAN”, *IEEE 2nd International Conference on 3G Mobile Communication Technologies*, 26-28 March 2001, pp. 47-52

[18] S. Maniatis, C. Grecas and I. Venieris, “End-to-End QoS Issues Over Next Generation Mobile Internet”, *IEEE Symposium on Communication and Vehicular Technology*, 2000, SVCT-2000, 19 Oct, 2000, pp. 150-154

[19] P. Newman, Netillion Inc. “In Search of the All-IP Mobile Network”, *IEEE Communication Magazine*, vol. 42, issue 12, Dec. 2004, pp. S3-S8

[20] G. Araniti, F. Calabro, A. Iera, A. Molinaro and S. Pulitano, “Differentiated Services QoS Issues in Next Generation Radio Access Network: a New Management Policy for Expedited Forwarding Per-Hop Behavior”, *IEEE Vehicular Technology Conference*, VTC 2004-Fall, vol. 4, 26-29 Sept. 2004, pp. 2693-2697

[21] S. Uskela, “All IP Architectures for Cellular Networks”, *2nd International Conference on 3G Mobile Communication Technologies*, 26-28 March 2001, pp. 180-185

[22] Jeong-Hyun Park, “Wireless Internet Access for Mobile Subscribers Based on GPRS/UMTS Network” *IEEE Communication Magazine*, vol. 40, issue 4, April 2002, pp. 38-39

[23] K. Daniel Wong and Vijay K. Varma, “Supporting Real-Time IP Multimedia Services in UMTS”, *IEEE Communication Magazine*, vol. 41, issue 11, Nov. 2003, pp. 148-155

[24] 3GPP, “Universal Mobile Telecommunication System (UMTS); QoS Concepts and Architecture”, *TS23.107V6*, March 2004

[25] R. Chakravorty, J. Cartwright and I. Pratt, “Practical Experience with TCP over GPRS”, *in IEEE GlobeCom*, Nov. 2002

[26] D. Schwab and R. Bunt, “Characterizing the use of a Campus Wireless Network”, *in IEEE INFOCOM*, March 2004

[27] X. Meng, S. Wong, Y. Yuan and S.Lu, “Characterizing Flows in Large Wireless Data Networks”, *in ACM Mobicom*, Sep 2004

[28] A. Balachandran, G. M. Voelker, P. Bahl and P. Venkat Rangan, “Characterizing user behavior and network performance in a public Wireless LAN”, *Sigmetrics Performance Evaluation. Review*, vol. 30. no. 1, 2002, pp. 195-205

[29] A. Adas and A. Mukherjee, “On Resource Management and QoS guarantees for long-range dependant traffic”, *In Proc IEEE INFOCOM*, 1995, pp. 779-787

[30] M. Parulekar and A. Makowski, “Tail Probabilities for a Multiplexer with self-similar input”, *In proc IEEE INFOCOM*, 1996, pp. 1452-1459

[31] I. Norros, “A Storage Model with self-similar input”, *Queueing System*, 16, 1994, pp. 387-396

[32] B. Tsybakov and N. D. Georganas, “Self-Similar traffic and upper bounds to buffer overflow in ATM queue”, *Performance Evaluation*, 36, 1998, pp. 57-80

[33] M. Zukerman et al, “Analytical Performance Evaluation of a Two Class DiffServ link”, *IEEE ICS*, 25-28 Nov. 2002, vol. 1, pp. 373-377

[34] J. M. Chung, Z. Quan, “Impact of Self-Similarity on Performance Evaluation in DiffServ Networks”, *IEEE MWSCAS*, 4-7 Aug. 2002, vol. 2, pp. 326-329

[35] C. F. Chou et al, “Low Latency and efficient packet scheduling for streaming applications”, *IEEE ICC*, 20-24 June, 2004, vol. 4, pp. 1963-1967

[36] A. Kos and B. Klepec, “Performance of VoIP applications in a simple Differentiated Services network architecture”, *IEEE EUROCON*, 4-7 July, 2001, vol. 1, pp. 214-217

[37] J. M. Chung and H. M. Soo, “Analysis of non preemptive priority queueing of MPLS networks with Bulk arrivals”, *IEEE MWSCAS*, 4-7 Aug. 2002, vol. 3. pp. 81-84

[38] Salil S. Kanhere and Harish Sethu, “Fair, Efficient and Low-Latency Packet Scheduling using Nested Deficit Round Robin”, *Proceedings of the IEEE Workshop on High Performance Switching and Routing (HSPR)*, May 2001

[39] N. F. MIR and A. Chien, “Simulation of Voice over MPLS communication networks”, *IEEE ICCS*, 25-28 Nov. 2002, vol. 1, pp. 389-393

[40] A. Krendzel, Y. Koucheryavy, J. Harju and S. Lopatin, “Traffic and QoS management in Wireless Multimedia Networks” COST 290:: *Wi-QoST*, Working group N3 http://www.cost290.org

[41] M. Jiang, M. Nikolic, S. Hardy and L. Trajkovic, “Impact of Self-Similarity on Wireless Data Network Performance”, *IEEE ICC*, 2001, vol. 2, pp. 477-481

[42] J. Ridoux, A. Nucci and D. Veitch, “Characterization of Wireless Traffic based on Semi-Experiments”, *Technical Report-LIP6*, December 2005

[43] Z. Sahinoglu and S. Tekinay, “On Multimedia Networks: Self-Similar Traffic and Network Performance”, *IEEE Communication Magazine*, vol. 37, issue 1, Jan. 1999, pp. 48-52

[44] I. Norros, “On the use of Fractional Brownian Motion in theory of connectionless networks”, *IEEE Journal on Selected Areas in Communications*, vol. 13. no. 6, August 1995, pp. 953-962

[45] P. Benko, G. Malicsko and A. Veres, “A Large-scale, passive analysis of end-to-end TCP Performances over GPRS”, *in IEEE INFOCOM*, March 2004

[46] M. Caglar, “A Long-Range Dependant Workload Model for Packet Data Traffic”, *Mathematics of Operations Research*, 29, 2004, pp. 92-105

[47] H. P. Schwefel, L. Lipsky, “Impact of aggregated self-similar ON/OFF traffic on delay in stationary queueing models (extended version)”, *Performance Evaluation*, 43, 2001, pp. 203-221

[48] I. Kaj, “Limiting fractal random processes in heavy-tailed systems”, In Fractals in Engineering, New Trends in Theory and Applications, Eds.J. Levy-Lehel, E. Lutton, *Springer-Verlag London*, 2005, pp. 199-218

[49] S.M. Ross, “Introduction to Probability Models” Academic Press, 1997.

[50] K.S. Trivedi, Probability and statistics with reliability, queueing, and computer science applications Wiley, New York, 2002.

[51] E. Cinlar, “Introduction to Stochastic Processes”, 1975, pp. 178

[52] R. Ben Ali, Y Lemieux and S. Pierre, “UMTS-to-IP QoS Mapping for Voice and Video Telephony Services, *IEEE Network* , vol. 19, issue 2, March/April 2005, pp. 26-32

[53] Y. Cheng, H, Jiang, W, Zhuang, Z. Niu and C. Lin, “Efficient Resource Allocation for China’s 3G/4G Wireless Networks, *IEEE Communication Magazine*, vol. 43, issue 1, Jan 2005, pp. 76-83

[54] W. Odom and M. J. Cavanaugh, “IP Telephony Self-Study Cisco DQoS Exam Certification Guide”, Cisco Press, 2004, pp. 3-314

APPENDIX

Table: 1 The States of the Markov Chain and Transition Probabilities

Initial State	**Reachable States** ($m=1,2,3,4$)	**Transition Probability**
$(i_1,i_2,i_3,i_4,a_1,s_1)$	$(i_1-k+1,i_2,i_3,i_4,a_m,s_1)$, $k=0,\dots,i_1$	$\int_0^\infty\int_0^t\int_{t-x}^\infty f_{S_1}(s)\,f_{S_1^k}(x)\,f_{T_{1m}}(t)\,ds\,dx\,dt$
	$(0,i_2-k,i_3,i_4,a_m,s_2)$, $k=0,\dots,i_2-1$	$\int_0^\infty\int_0^t\int_{t-x}^\infty f_{S_2}(s)\,f_{S_1^{i_1+1}+S_2^k}(x)\,f_{T_{1m}}(t)\,ds\,dx\,dt$
	$(0,0,i_3-k,i_4,a_m,s_3)$ $k=0,1,\dots\dots.i_3-1$	$\int_0^\infty\int_0^t\int_{t-x}^\infty f_{S_3}(s)\,f_{S_1^{i_1+1}+S_2^{i_2}+S_3^k}(x)\,f_{T_{1m}}(t)\,ds\,dx\,dt$
	$(0,0,0,i_4-k,a_m,s_4)$ $k=0,1,\dots\dots\dots\dots.i_4-1$	$\int_0^\infty\int_0^t\int_{t-x}^\infty f_{S_4}(s)\,f_{S_1^{i_1+1}+S_2^{i_2}+S_3^{i_3}+S_4^k}(x)\,f_{T_{1m}}(t)\,ds\,dx\,dt$
$(i_1,i_2,i_3,i_4,a_1,s_2)$	$(i_1-k+1,i_2-1,i_3,i_4,a_m,s_1)$, $k=0,\dots,i_1$	$\int_0^\infty\int_0^t\int_{t-x}^\infty f_{S_1}(s)\,f_{S_1^k+S_2^1}(x)\,f_{T_{1m}}(t)\,ds\,dx\,dt$
	$(i_1+1,i_2,i_3,i_4,a_m,s_2)$	$\int_0^\infty\int_t^\infty f_{S_2}(s)\,f_{T_{1m}}(t)\,ds\,dt$
	$(0,i_2-k,i_3,i_4,a_m,s_2)$ $i_2=2,3,\dots$ and $k=1,\dots,i_2-1$	$\int_0^\infty\int_0^t\int_{t-x}^\infty f_{S_2}(s)\,f_{S_1^{i_1+1}+S_2^k}(x)\,f_{T_{1m}}(t)\,ds\,dx\,dt$
	$(0,0,i_3-k,i_4,a_m,s_3)$ $k=0,1,\dots\dots\dots.i_3-1$	$\int_0^\infty\int_0^t\int_{t-x}^\infty f_{S_3}(s)\,f_{S_1^{i_1+1}+S_2^{i_2}+S_3^k}(x)\,f_{T_{1m}}(t)\,ds\,dx\,dt$

	$(0,0,0,i_4-k,a_m,s_4)$ $k=0,1,2,\ldots\ldots\ldots\ldots i_4-1$	$\int_0^\infty \int_0^t \int_{t-x}^\infty f_{S_4}(s)\, f_{S_1^{i_1+1}+S_2^{i_2}+S_3^{i_3}+S_4^k}(x)\, f_{T_{1m}}(t)dsdxdt$
$(i_1,i_2,i_3,i_4,a_1,s_3)$	$(i_1-k+1,i_2,i_3-1,i_4,a_m,s_1)$ $k=0,1,\ldots\ldots\ldots\ldots i_1$	$\int_0^\infty \int_0^t \int_{t-x}^\infty f_{S_1}(s)\, f_{S_1^k+S_3^1}(x)\, f_{T_{1m}}(t)dsdxdt$
	$(0,i_2-k,i_3-1,i_4,a_m,s_2)$ $k=0,1,\ldots\ldots\ldots\ldots i_2-1$	$\int_0^\infty \int_0^t \int_{t-x}^\infty f_{S_2}(s)\, f_{S_1^{i_1+1}+S_2^k+S_3^1}(x)\, f_{T_{1m}}(t)dsdxdt$
	$(i_1+1,i_2,i_3,i_4,a_m,s_3)$	$\int_0^\infty \int_t^\infty f_{S_3}(s)\, f_{T_{1m}}(t)\, ds\, dt$
	$(0,0,i_3-k,i_4,a_m,s_3)$ $i_3=2,3,\ldots\ldots \quad k=1,\ldots,i_3-1$	$\int_0^\infty \int_0^t \int_{t-x}^\infty f_{S_3}(s)\, f_{S_1^{i_1+1}+S_2^{i_2}+S_3^k}(x)\, f_{T_{1m}}(t)\, dsdxdt$
	$(0,0,0,i_4-k,a_m,s_4)$ $k=0,1,\ldots\ldots\ldots\ldots\ldots i_4-1$	$\int_0^\infty \int_0^t \int_{t-x}^\infty f_{S_4}(s)\, f_{S_1^{i_1+1}+S_2^{i_2}+S_3^{i_3}+S_4^k}(x)\, f_{T_{1m}}(t)dsdxdt$
$(i_1,i_2,i_3,i_4,a_1,s_4)$	$(i_1-k+1,i_2,i_3,i_4-1,a_m,s_1)$ $k=0,1,\ldots\ldots\ldots\ldots i_1$	$\int_0^\infty \int_0^t \int_{t-x}^\infty f_{S_1}(s)\, f_{S_1^k+S_4^1}(x)\, f_{T_{1m}}(t)dsdxdt$
	$(0,i_2-k,i_3,i_4-1,a_m,s_2)$ $k=0,1,\ldots\ldots\ldots\ldots i_2-1$	$\int_0^\infty \int_0^t \int_{t-x}^\infty f_{S_2}(s)\, f_{S_1^{i_1+1}+S_2^k+S_4^1}(x)\, f_{T_{1m}}(t)dsdxdt$
	$(0,0,i_3-k,i_4-1,a_m,s_3)$ $k=0,1,\ldots\ldots\ldots\ldots i_3-1$	$\int_0^\infty \int_0^t \int_{t-x}^\infty f_{S_3}(s)\, f_{S_1^{i_1+1}+S_2^{i_2}+S_3^k+S_4^1}(x)\, f_{T_{1m}}(t)dsdxdt$
	$(i_1+1,i_2,i_3,i_4,a_m,s_4)$	$\int_0^\infty \int_t^\infty f_{S_4}(s)\, f_{T_{1m}}(t)dsdt$
	$(0,0,0,i_4-k,a_m,s_4)$ $k=1,2,\ldots i_4-1$ $i_4=2,3,\ldots\ldots\ldots\ldots$	$\int_0^\infty \int_0^t \int_{t-x}^\infty f_{S_4}(s)\, f_{S_1^{i_1+1}+S_2^{i_2}+S_3^{i_3}+S_4^k}(x)\, f_{T_{1m}}(t)dsdxdt$