1 Modeling Variability

Variability in the rate of arrival processes has been a characteristic of data communication between or among computers from the earliest days of networking. Back in the time when remote computers were accessed through terminals, communication lines carried character traffic, and computers were not linked to each other through a network, the rate of arrival of characters did not fluctuate a great deal. The primary cause of variability was human: users paused for different lengths of time to think between typing intervals.

The introduction of networking technology and its refinement has, however, directly led to an increase in the variability of arrival processes. First, the appearance of real computer networks, file transfer protocols, and computer mail resulted in more variability, as users could now instruct a system to send a message to a remote location, or copy data from one location to another. Measurement studies of wide-area network traffic revealed that the traffic was bursty [47]. Next, the advent of higher-speed local-area networks and the simultaneous development of the personal computer workstation [57,90] opened fascinating new avenues for data communication. Remote file systems, distributed systems, more sophisticated protocols, including the Remote Procedure Call protocol (RPC), and distributed applications became the norm. All of these advances contributed to increasing variability since faster protocol allowed for transmission of more data in a shorter time, while many distributed-system operations generated successive transmission requests within a wide range of times. Studies showed that local-area traffic was also bursty [80], but this was not a matter of concern, because local-network traffic was extremely light and there were no noticeable queuing delays.

Variability increased even more dramatically when diskless workstations made their debut. With diskless workstations, virtual-memory traffic had to be moved as quickly as possible across the network. A few keystrokes could produce either a quick burst of up to several hundred Kbytes of data across the network when a new program was invoked, or only a minimal data representing, for instance, characters transmitted to a remote computer.

While the networking research community began to experience congestion and to think about ways to cope with it, telecommunication researchers, who had long been studying telephone voice channels, had already well understood the effects of bursty arrival processes to queues. Consider, for instance, the following two cases. In the first case, we have a single-queue system whose input is a Poisson process with arrival rate λ; in the second case, the input process is batch Poisson, i.e. arrivals occur, according to a

Poisson process, in batches of mean size B. If the arrival rate of the batch-Poisson process is chosen to be λ/B, the mean arrival rate is the same in each of the two cases. The average times spent in the systems, according to Little's law, will be proportional to the number of customers in the systems. The mean number of customers in the system for a Poisson process is $N_P = \frac{\rho}{1-\rho}$. The mean number of customers in the system in the batch-Poisson case can be calculated easily [46, problem 4.3]:

$$N_B = \frac{1}{2}\left(1 + E(B) + \frac{VAR(B)}{E(B)}\right) N_P \ .$$

The quantity between parentheses is always greater than 2 (it equals 2 when all batches are of size 1, i.e. when the batch-Poisson process reduces to a Poisson process), and hence the average waiting time of customers arriving according to a Poisson process is shorter than the waiting time of customers arriving according to a (more bursty) batch-Poisson process with the same mean arrival rate.

Despite its profound effects on computer network traffic, queuing, and congestion, variability has been studied mostly in terms of analytical or simulation models. In this dissertation, we will consider measured packet-arrival processes, study their variability based on indices of dispersion, and propose models that will characterize this variability.

II. Central Problem of Dissertation

This thesis examines the problem of characterizing the variability of packet-arrival processes produced by individual workstations in a local-area network. The study is based on measurements and emphasizes understanding and modeling real data. These data are packet-arrival processes extracted from the arrival times of aggregate traffic produced by engineering workstations on an Ethernet local-area network. The term ``processes'' refers to realization of stochastic processes along a time axis. Technically, they are marked point processes as each point has several attributes, such as a packet length, a protocol type, a source and a destination network addresses, and so on. Often we will derive the properties of the point processes by operating on time series of interarrival times—the series of intervals between successive packet arrivals.

Of all the features of complex phenomena such as packet arrival processes—a result of the complexity of the systems that produce them—we will focus on variability. Often referred to as burstiness, variability can be informally described as producing a string of abrupt transitions from low to high arrival rates and vice versa. We found no satisfactory quantitative definition of variability in the computer science literature, but since variability can be related to the relationship among successive arrivals, we adopted a definition from the theory of point processes based on indices of dispersion.

We have chosen to study packet arrival processes generated by engineering workstations because these workstation constitute an important example of distributed system whose network usage patterns are not completely understood. We conducted the measurements on a local-area network, and not on a wide-area network, in order to keep the measurement effort manageable, but also because arrival-process variability is more evident in a local-area environment where there are fewer queuing interactions and less traffic smoothing produced by the superposition of traffic generated by

machines communicating concurrently. There are other types of interesting systems that produce network traffic worth modeling; transaction traffic generated by distributed database systems comes to mind; however, given the intense research effort [38] in the design of the new generation of workstations that will handle continuous media applications in addition to current user interfaces, we felt that models (even if only models of variability) of current workstation traffic could have a direct impact on the workstation design.

Though a major portion of the dissertation is devoted to the analysis of the first- and second-order statistical properties of interarrival-time and packet-count series, which reveal the structure of the underlying point processes produced by individual workstations, we will also define models of traffic. These models are not intended to reproduce the traffic exactly as generated by the workstations. Instead, the models will focus on the notion of variability of packet arrival processes. Using indices of dispersion, which express the autocorrelation structure of a process, we can quantify data variability; we can then incorporate into mathematical models the second-order stochastic parameters that represent dispersion, and derive indices of dispersion for the models. Fitting is done so that the models' variability closely matches that of the data. This approach produces what we have defined as ``models of variability''. The gist of the argument is that the autocorrelation structure of a point process depends on the lengths of sequences of observations smaller than the mean, on the lengths of sequences larger than the mean, and the way the two types of sequences alternate. Hence, the form of the solution that we seek to the problem of characterizing the variability of packet arrival processes is one in which models will reproduce the correlation structure the the measured point processes, not necessarily the exact marginal distribution of their interarrival times.

III. Thesis Roadmap

Chapter 2 deals with the measurement methodology, the data collection, the preparation of the data for analysis, and the initial data analysis. The data was collected on a large Ethernet, which we will denominate ``Engineering network'', at the Sun Microsystems headquarters in April 1989. The network consisted of about 130 machines, divided among diskless workstations and machines with local disks. The workstations were served by six major file servers. In the thesis we study arrival processes that occurred approximately between 8:15 am and 6:30 pm on a single day. Based on observing the measured network over a period of several days, we are confident that the chosen day is typical. The data was packaged into traces, which, stored on tapes, consist of packet headers and timing information.

The data collection was preceded by several months of study for the preparation and design of the measurement system. The Performance Evaluation group of Sun let us use the fastest machine available at the time. We knew from our previous measurement project (described in Appendix A) that data analysis would greatly benefit from the availability of an accurate clock with which to timestamp packet arrivals. As the resolution of the machine clock was only 10 ms, a serious limitation, we employed an external timing device, described in Section 2.III, whose resolution was 0.5 microseconds.

Data analysis required reading the traces innumerable times. This would not have been practical if we had kept the data on tapes, as tape-drive access times are too

slow. Since we could not monopolize many hundred Mbytes of disk space for our own use, we were forced to compress the trace data; this is described in Section 2.VI. While compressing the data, we also transformed the timing information, which was recorded after each packet had been fully received, to the times when the packet transmissions started. This is possible because most of the packet-delay components are deterministic. The random components of the delay, most importantly the interrupt time latency, contribute to the measurement errors.

During the timing conversion, we noticed that for a sizable portion of the packet arrivals, the beginning of the transmission time should have occurred before the transmission of the previous packet had completed. This, on a single-channel, baseband network, is impossible. Hence, the situation must have been created by measurement errors. In Section 2.VII, we analyze several cases of errors in this class and, in all cases, we show that errors were generated by pairs of packets that were transmitted with the minimum interpacket time admissible on the Ethernet, 16 microseconds. Our analysis attributes the error source to the tracing machine's internal bus, which was shared by the CPU, the network interface, and the timing device.

At the end of Chapter 2, in Section VIII, we describe some of the characteristics of the aggregate traffic of the Sun Engineering network. We discuss the network utilization, the source-destination traffic patterns, the interarrival-time probability density, and the distribution of packet lengths. This data analysis should be compared to the analysis of older measurement that we took at Berkeley in an environment of diskless machines. (The major results of this study are reported in Appendix A.)

Chapter 3 deals with the stochastic analysis of the traffic produced by individual workstations. The point processes produced by the workstations are packet arrivals at the Ethernet data-link layer. As such, they are different from the processes representing arrivals of messages produced by users. Consider for instance a typical 8-Kbyte user read operation: the single user request is translated by the data-link layer into a sequence of six fragments. In addition, data-link-layer packet arrivals are affected by network contention: were source and destination machines the only communicating machines the arrival patterns would be different than those extracted from the aggregate traffic. However, based on the analyses in Section 3.IX, we are confident that the errors introduced by network contention are small.

One of the most important questions is whether arrival processes (or, equivalently, interarrival time series) are stationary. Our packet-arrival processes appear, even to summary investigation, clearly nonstationary. In Section 3.V we discuss whether standard statistical techniques for removing nonstationarity can be applied to interarrival time series. We conclude that those technique, all of which amount to filtering the data so as to remove the lowest frequencies, would disrupt the data appearance and would prevent us from recognizing arrival process features and associating them to system parameters.

We recognize that the question of whether a time series is stationary is intimately related to the time scale under consideration. There are at least two time scales of interest in arrival processes that are generated by protocols under user-driven workloads: a coarse one that corresponds, for instance, to user's busy and idle intervals, and a fine one that captures protocol time constants. The key observation is that user-generated events occur infrequently when compared to protocol-generated ones. Hence, a

coarse time scale that represents macroscopic changes in the packet arrival patterns is more appropriate for characterizing user activity. If instead the focus is on protocol behavior, a finer time scale, ranging from a fraction of a millisecond to several hundred milliseconds, is more desirable. In Section 3.V we observe 1) that we do not have enough data to attempt a user characterization, and 2) that a fine time scale is more important to the queuing dynamics of packet arrival processes in interactive systems in which response time is the key performance index. Thus, we decide to restrict the analysis to fine time scales, and we remove the major nonstationary components by considering only interarrival times shorter than a fixed threshold, which we select in Section 3.VI. (Unfortunately, this truncation operation does not fully eliminate nonstationarity.)

In Chapter 3, we also discuss network utilization and histograms of interarrival times, which are related to marginal probability densities. Using the properties of our marked processes (for instance packet lengths, and source/destination addresses), we partition interarrival times into components. The partition shows that short packets produce long interarrival times and, vice versa, that long packets produce short interarrival times. We then study the autocorrelation structure of our processes, and, in the last section, examine whether processes produced by independent machines are statistically related because of network contention and concurrent access to file servers. The result is that unrelated processes are also uncorrelated.

We also carried out the spectral analysis of interarrival times, but did not include it in the dissertation. Peaks in the power spectra of intervals are expressed in terms of lags. Charnock in [13] addresses the problem of converting the lag references to times. Aside from the complication introduced by this procedure, the frequency-domain analysis revealed only, and was dominated by, the effects introduced by packet fragmentation; no other relevant phenomenon was discernible.

In Chapter 4 we define the indices of dispersion for intervals and for counts, and examine their properties. They are series of coefficients related to the autocorrelation coefficients of intervals and counts. These indices are the pivotal elements of our variability characterization. We seek models that approximate well the indices of dispersion, or, equivalently, the series of autocorrelation coefficients.

In Section 4.IV we give an example of a model (though a somewhat artificial one) in which parameters are fitted using the index of dispersion for counts. The counts produced by the models may not be similar to those of the data (they are related to them because the mean and the variance are fitted), but, by matching the correlation coefficients, the model generates sequences in which groups of counts larger than the mean and groups of counts smaller than the mean have a structure that follows the structure of similar groups in the data. The unproven assumption in this dissertation is that if data arrivals and the arrivals produced by the model are submitted to the same server of a queuing systems, they will produce comparable queuing effects.

Chapter 5 is devoted to the definition of the models of variability. These are models of interarrival times based on what we call a ``generalized two-state semi-Markov model'': a semi-Markov model in which intervals in each state are correlated, though intervals belonging to different states are independent. Our modeling objective is to reproduce the variability rather than to insure that the queuing behavior of the models approximates to some degree the queuing behavior of the data. We believe that there

is a close relationship between the two, but the extent to which this is true will have to be explored with further research. (This question is discussed in the Future-Work section in Chapter 6.)

The distribution of packet lengths studied in Chapter 2 indicated a sharp repartition of packet shorter than 200 bytes (short packets) on one side and those longer than 500 bytes (long packet) on the other side; there are virtually no occurrences of lengths between 200 and 500 bytes. To build the models, we first show that the lengths of short packets and the lengths of long packets constitute a two-state Markov chain: one state for short packets, the other for long packets. Second, we associate to the states of the Markov chain the probability density functions (component functions) of short and long interarrival times. Third, we derive the index of dispersion for intervals for the generalized semi-Markov model in terms of the autocovariances of the interarrival densities of the two states. A nice, simplifying feature of a semi-Markov model, which is also valid for the generalized model, is that its autocorrelation coefficients depend only on the first and second-order properties of the component density functions.

We discover that the model does not fit nonstationary data and that, to obtain a good match, we have to limit its range of applicability to lags up to 30 and to interarrival times up to about 500 ms. However, these are the intervals over which the variability of arrival processes, produced and consumed by fast, interactive queuing systems, is relevant.

This concludes the thesis outline; in the reminder of this chapter we will examine some of the published works that are related to several aspects of the dissertation.

IV. Literature Review

In this section we will briefly review the major works that are related to the dissertation. We will consider two classes of studies: network measurements and analysis studies, and packet arrival process modeling with emphasis on arrival-variability modeling.

A. Network Measurements and Their Analysis

One of the first measurement studies of local-area network traffic was performed in 1979 at Xerox PARC by Shoch and Hupp [80]. Shoch and Hupp suggested that the Ethernet under its normal workload, produced by users of Alto [90] computers accessing file servers, is lightly loaded and the host interfaces, and not the network, are the protocols' performance bottlenecks.

More recently, Boggs et al. [9] have performed measurements of artificial traffic to explore the limits of Ethernet behavior.

D.R. Cheriton and C.L Williamson [15] measured the traffic properties of an Ethernet at Stanford University that connects about 50 diskless Sun workstations running the V operating system [16]. The authors focused on the performance of the VMTP request-response protocol [14], a general-purpose protocol used for all network communication in V. They did not quantify the network utilization (which was rather low) since they were mainly concerned with the interactive characteristics of their system, such as the message transaction rate, the duration of transactions, and VMTP's retransmission behavior. The measured V systems did not have virtual memory; as a result, most of the traffic was

generated by file accesses and V's distributed-system services such as the naming service. They express the opinion that the current generation of protocols is not suitable for future high performance workstations.

E.D. Lazowska et al. [49] measured a network of diskless Sun-2s and their file servers in the context of a modeling study in 1986. Surprisingly, they found that their workstations, equipped with 2 Mbytes of physical memory, display a 4:1 file access to paging ratio. Such a finding could be explained in part by the fact that UNIX applications were smaller at that time.

In a study on network file-system caching in the Sprite operating system, M. Nelson et al. [59] concluded that even without client caching, Sun-2 machines require network bandwidths on the order of 80 Kb/s and Sun-3s on the order of 250 Kb/s; in the Berkeley study [33] we observed 1 Mb/s between Sun-3s with NFS client-side caching. The difference can be explained by the fact that they ran benchmarks consisting of compilations, sorting tasks, and text formatting jobs rather than simulating the full workload generated by a user running processes in a window environment.

P.J. Leach et al. have designed an integrated distributed system where diskless workstations and file servers communicate via a 12 Mb/s baseband token-passing ring network [50]. Preliminary performance measurements indicate that the average network utilization is low: less than 1/12 of the total network bandwidth was used. However, paging alone consumes ten percent of the network bandwidth when running artificial workload tests. Since in the described implementations the maximum packet size was 550 bytes–their network interfaces allow packets as large as 2048 bytes–higher utilization will be achieved by future, more efficient implementations.

A number of measurement studies point out the presence of high burstiness in packet data, but detailed models that incorporate this variability explicitly are not given [33,54]. Measurement studies have been recently reviewed and classified by Pawlita [65]; we refer the reader to Pawlita's comprehensive taxonomy for additional references.

Newer measurement studies [51,93] have used high-resolution clocks to obtain timing information with which it is possible to estimate statistical distributions with small bias.

Wide area traffic was measured recently by Caceres [12] and Heimlich [36]. This studies are interesting though only minimal statistical analyses are presented.

B. Arrival Process Modeling

Here we review some contributions in the area of the analysis of variability in arrival processes. The key observation is that a number of models have been developed to deal with data variability; however, most models have not been applied to computer communication traffic data or to real data at all.

In the 60's a number of statistical studies emphasized the notion that many interesting point processes cannot be described accurately as Poisson processes [6,52]. These efforts culminated in the publication of a path-breaking book by Cox and Lewis [19], whose characterization of the second-order properties of point processes is still used today in the recently revived interest in the study of data variability spurred by growing use of communication networks. For instance, Sriram and Whitt [84] propose to use the index of dispersion for intervals (IDI) for characterizing the superposition of voice and

data arrival processes, and in a recent contribution Fendick and Whitt [26] propose to use a scaled version of the variance-time curve (the index of dispersion for work) as a measurement to describe the variability of the traffic offered to a queue.

Ramaswami and Latouche [74] propose a unified stochastic model for the arrival of voice packets to a packet switch. They develop a point-process model that takes into account the statistical fluctuations of packet generation of individual calls. For this, they also resort to the use of index of dispension curves.

Heffes and Lucantoni [35] use a Markov Modulated Poisson Process (MMPP) to model the dependencies present in voice and data sources. They obtain the moments of the voice and data delay distributions as well as the queue length distribution. Numerical results for the tails of the voice packet delay distribution show the dramatic effect of traffic variability and correlations on performance.

Jain and Routhier [42] propose clustering techniques to capture dependencies in packet streams and apply these techniques to packet data collected on a ring network at MIT [24]. Trains are sequences of packets that are considered as a unit. They are identified by a single parameter: the *intercar gap*. If the interpacket time is larger that the intercar gap, then the next packet will start a new train, otherwise the next packet is part of the current train. The authors explored several variations on the theme, but seem to ignore any physical interpretation: their mean intertrain time is 23.8 seconds! They define their trains arbitrarily and give no suggestion as to how one should chose the maximum intercar gap.

Jain and Routhier's approach has been carried on to a more theoretical basis by Fontana and Guerrero, who propose a Markov model extended to include packet trains [29].

Song and Landweber propose to modify the protocol layers on host machines in order to optimize the transmission and receipt of trains of packets [83]. The proposed packet-train mechanism takes advantage of the application of similar protocol processing functions to all packets associated with a single bulk transfer. They studied the performance of their scheme with simulation and tried different train sizes, but offer no insight on how to choose the sizes of trains, nor they seem to rely on any particular train model.

Fendick et al. [25] in a theoretical study associate the burstiness of data in communication networks to various correlations in the data.

The variability characteristics of a process generated as the superposition of many sources has been studied by Albin [1], by Sriram and Whitt [84], and Heffes and Lucantoni [35].

Neuts [60, page 336] is interested in developing mathematically tractable descriptors of the physical behavior of point processes. Among such descriptors are the moments of the counting process, the peakedness functional, the power spectral density of a square wave (random telegraph wave) associated with the point process, and the caudal characteristic curve, which was introduced in [61]. The (exponential) peakedness function of a stationary point process is the ratio of the variance and the mean of the number of busy servers at an arbitrary time in an infinite server queue with exponential service times and the given point process as input.

Cruz [22] develops a calculus for obtaining bounds on delay based on input data streams that satisfy ``burstiness constraints''. A data stream is said to satisfy a burstiness constraint if the amount of data from the stream contained in any interval of time is less than a value that depends on the length of the intervals.

In the context of operating-system and networking research aimed at providing support for continuous media, Anderson et al. [3] define a model for traffic sources based on Cruz's thesis work. Their abstraction has the following parameters: a maximum message size of S_{max} bytes, a maximum message rate of R_{max} messages/second, and a maximum burst size of B_{max} messages. The long-term data rate is $S_{max}R_{max}$ bytes/second. In any interval of length t, the number of messages arriving at a network interface may not exceed $B_{max} + tR_{max}$. The burst parameter allows short-term violations of this rate constraint. An analogous, simple model is proposed by Ferrari and Verma [27].

Leland's study of LAN traffic [51] should be emphasized as it is similar in methodology and basic results to our measurements and analyses in Chapters 2 and 3. Leland develops the aggregate data analysis more than we do, and, although he mentions possible models for various data characteristics, he does not develop a traffic model. He associates explicitly the notion of burstiness to indices of dispersion.