Communication Models for Third Party Load Frequency Control

Sudipto Bhowmik, Kevin Tomsovic, *Senior Member, IEEE*, and Anjan Bose, *Fellow, IEEE*

***Abstract*—With deregulation of the power generation sector, the necessity for an enhanced and open communication infrastructure to support an increasing variety of ancillary services is apparent. A duplex and distributed communication system seems to be the most suitable solution to meet and ensure good quality of these services. Parameters needed and additional limits introduced by this new communication topology must be investigated and defined. This paper focuses on the communication network requirements for a third party load frequency control service. Data communication models are proposed based on queuing theory. Simulation is performed to model the effects of certain types of signal delays on this ancillary service.**

***Index Terms*—G/G/1 queues, load frequency control, network delays, queueing theory.**

I. Introduction

INCREASED competition and the need for varied ancillary services for the power generation sector have led to the desire for a more open, adaptable and distributed communication network. The importance of such a communication system is underlined in the recent NERC Policy 10 [1]. For example, one of NERC's requirements is real-time voice and data communication that every supplier must maintain with the operating authority at the control center. A new communication infrastructure is fast becoming an exigent need not only to meet NERC's requirements but also for independent generation companies to offer third party services, such as load frequency control (LFC), at low cost. An improved communication infrastructure is important for the Independent System Operator (ISO), as they are responsible for the monitoring of all network components under their jurisdiction. For load following, communication may need to be provided, both, for a traditional central control offered as an ancillary service and also in the case of third party or bilateral contracts between generators and consumers. Several of the system operators are already expediting the migration to a distributed network, namely the internet, in order to meet this requirement.

Load following and frequency control are traditionally provided through automatic generation control (AGC). The AGC signals are sent via dedicated communication channels, which are the responsibility of the large utilities. Backup, in case of channel failure, was provided by voice communication via telephone lines. The new infrastructure should have redundant links to guarantee fault tolerance in case of link failures, as the penalty for not meeting the required generation profile can be steep. This is an important factor for migrating to a distributed infrastructure as it inherently offers this redundancy. In addition, bilateral contract opportunities place a dependency on communication, for meeting customer load with adequate quality of service. It will be shown that the third party generators are highly susceptible to delays in the control signals.

Traditionally, communication network analysis for parameters such as throughput and delays are performed using queuing theory. The models are largely based on an exponential arrival rate as it allows several simplifications to quantify the waiting time in queues. Contemporary research data of the internet suggest a more statistically similar distribution of the major internet traffic. Empirical research has been performed to quantify the internet traffic and to model them based on these premises [2]. We propose queueing models based on a network that may be included in the load following system model. These models are based on constant packet length instead of exponentially distributed packet length as is typically done. Scenarios considered, include failure and recovery of queuing servers and servers under a denial of service attack. Of late, greater emphasis has been placed on ensuring that the communication system remains robust in the face of malicious attacks. Most likely, physical network security will remain the weakest point for large scale damage, but one must understand the effects of other software related malicious attacks.

Recently, the possibility of allowing a bilateral market for the provision of load following and frequency control services has arisen provided there exists an appropriate communication channel [3]. For operation of the load frequency control a certain number of generating units will receive a signal input, in the form of a packet of data, as a raise or lower power output. This signal could originate from either the center responsible for the AGC or from the customer in the bilateral market.

The models suggested here include the data delays in the network layer but do not give an indication of the additional delays introduced due to routing, retransmission and other applications. Still, they offer good insight for basing further analysis to quantify requisite parameters for higher layer services and for further studies of robust controls. Based on these models, appropriate higher level protocols can be determined and tuned to meet high reliability requirements in the face of the heterogeneous nature of the load following entities. Simulations for the load following model and that of the network model are effectively disjoint as the sending of the signal is deterministic.

Manuscript received March 8, 2003. This work was supported in part under the EPRI/Dod Complex Interactive Networks Initiative and in part by the CERTS program at the Department of Energy.

The authors are with the School of Electrical Engineering and Computer Science, Washington State University, Pullman, WA 99164 USA (e-mail: sbhowmik@eecs.wsu.edu; tomsovic@eecs.wsu.edu; bose@wsu.edu).

Digital Object Identifier 10.1109/TPWRS.2003.818700

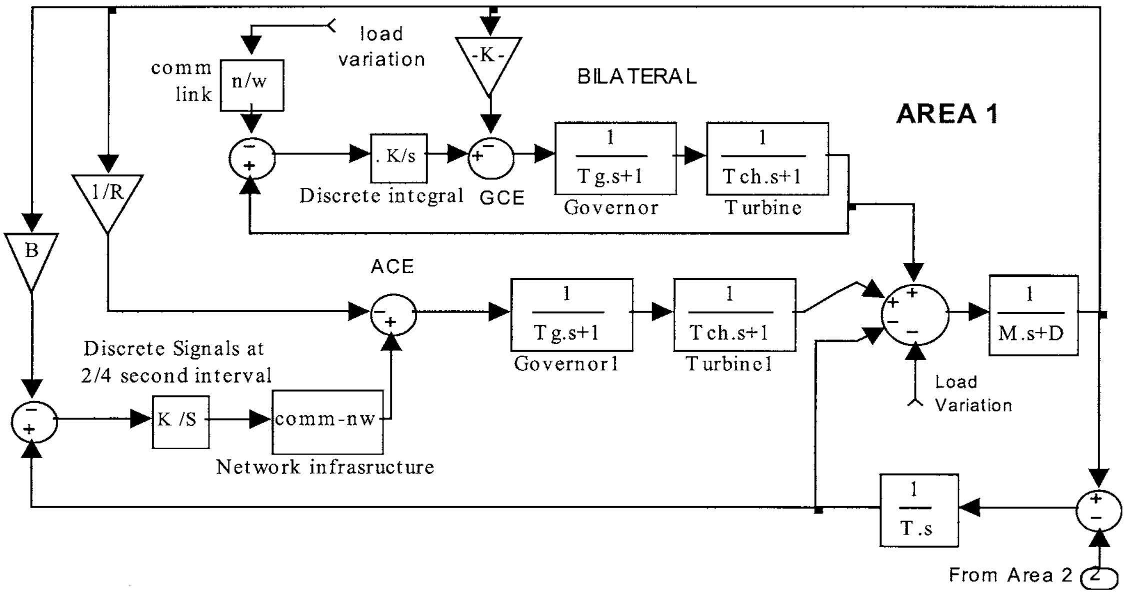


Fig. 1. Block diagram for AGC with bilateral contract.

II. Effect of Signal Delays on Load Following

Before introducing the various models, some effects of signal delays on the load following model are shown. Simulations are performed on three different systems. The models all consist of three control areas (CA) interconnected together via tie lines with sets of 9, 17, and 50 generators respectively. Here, CA 1 is considered a small area, CA 2 a medium area and CA 3 a large area to represent several possible different scenarios. Simulations were carried out with the generators configured to accept the classical AGC signals, a pure bilateral configuration for third party services [3], and a mixture of the two. Fig. 1 illustrates the bilateral and the classical model used for the simulations. It depicts a section of a control area. Standard simplified models for the prime mover and governor were used [4].

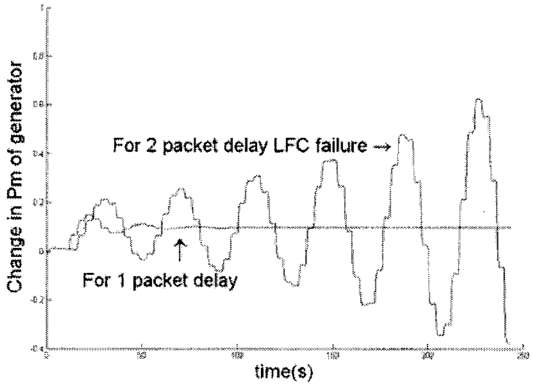
The area control error (ACE) and generator control error (GCE) signals were sent every 4 seconds as is typical in the US. Participation among the generators involved in the AGC service, were divided equally. In the case of a mixture of AGC and bilateral contracts, higher participation was assigned to the bilateral contracts. To model the delays, simulation based on constant packet delays as well as random delays, involving both individual generators as well as that induced at the source itself, (disseminator of AGC signals) were conducted. The constant delays denote a heavily congested network or a denial of service type attack at the respective site. The random delays denote Byzantine failures as well as malicious attacks. For the load variation, a step load increase is used. Simulations were done using Matlab with delays introduced. More details can be found in [5].

Results are summarized in Table I. In the simplified models used, failures lead to instabilities in the system. In the actual system, of course, protection and control logic would prevent such response. Still, the instabilities indicate LFC problems. Figs. 2(a) and (b) show that the bilateral generator is susceptible to a delay of only 2 packets. While this value is, of course, system dependent and also contingent on the generator parameters, clearly, it is of utmost importance for the bilateral contract entity to utilize a low latency communication channel with fault tolerance built into the control scheme. A mechanism for detecting old packets can be implemented by time stamping each packet and maintaining synchronized clocks (preferably virtual) for the successful utilization of the time stamps. In addition, some redundancy in the communication links is advisable.

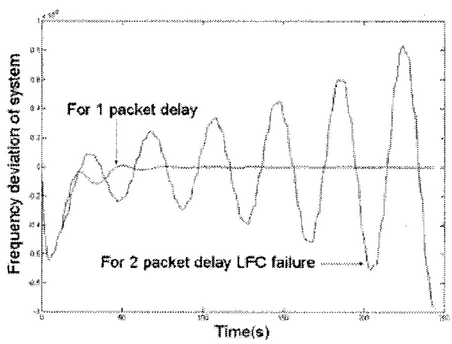
Centralized load following ancillary service is generally robust to signal delays, if the delays are present in a minority of the

TABLE I
Summary of Simulations

Scenario	**Centralized AGC**	**Third Party**	**Mixed AGC and Third Party**
Fixed delay	Failure of AGC for 15 packet delay in single generator. Failure at 3 packet delay in all generators.	Failure at 2 packet delay in any generator.	Failure at 2 packet delay in a bilateral generator. Failure at 3 packet delays in all AGC generators, or a 7-8 packet delay in majority of the generators. Delays tend to degrade system response.
Random delay	Failure at certain situations with random delay in all generators. No adverse affects from random delay in single generator.	Fail to meet customer demand and may become unstable.	System not adversely affected. Bilateral units may not meet the contractual schedule.
Both fixed and random delays	--	--	System may become unstable for short delays.

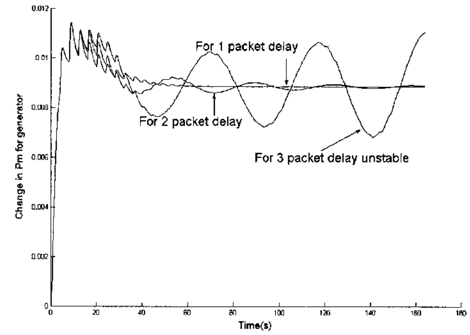


(a)

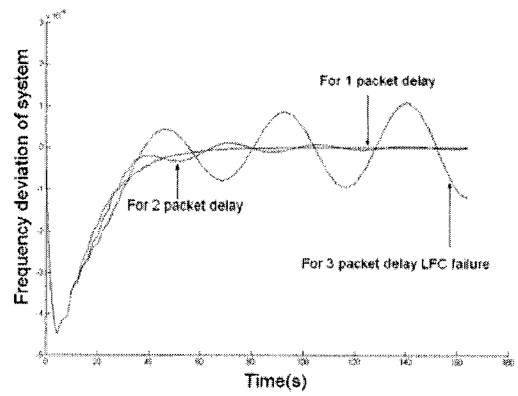


(b)

Fig. 2. (a) Power deviation in generator with bilateral contract given constant packet delays. (b) System frequency response with bilateral contracts given constant packet delays.



(a)



(b)

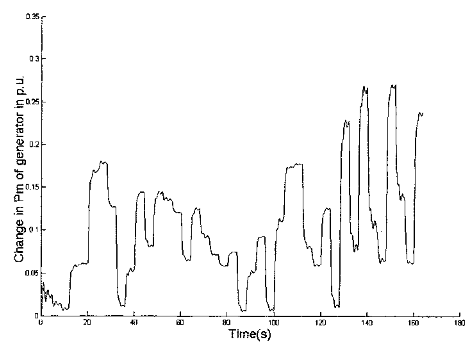
Fig. 3. (a) Power deviation for generator in centralized AGC given constant packet delays. (b) System frequency response for centralized AGC system given constant packet delays.

communication channels. However, if a majority of the channels exhibit delays in communication then the system response degrades. This is illustrated in Figs. 3(a) and (b), which show the system failing to operate correctly in the case of a 3 packet delay occurring in at least 66% of the AGC channels. This means the source of the AGC signals (presumably the ISO) is most susceptible to denial of service attacks or network congestion. As the controlling authority is typically the center for most data (both sending and receiving), this becomes a serious concern and appropriate measures must be taken to decrease the possibility of any such occurrence.

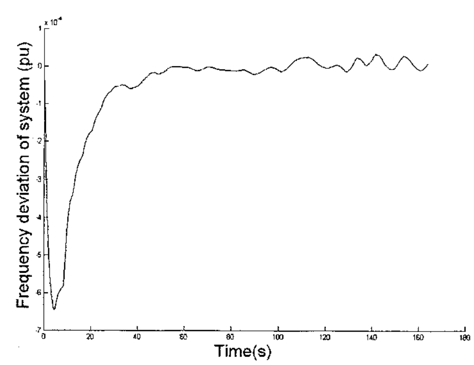
Fig. 4 shows that for random delays the bilateral contract cannot meet the contract with the customer, although the effect on the system is not very large Fig. 4(b). This failure to fulfill the contract may prove detrimental to the third party entities but at least does not cause system wide problems. Random delays can be introduced by Byzantine failures of the communication channel or by malicious parties.

III. Network Delay Models for LFC

Section II has demonstrated the importance of delays and random signals on the LFC. To analyze the network delays, queuing theory models are now introduced that mostly focus on packet delays in the network layer. The delay models suggested act as a way to ascertain whether a given network structure introduces delays that lead to unacceptable performance in the system. Thus, we are proposing to incorporate the appropriate communication delay model in any detailed simulations of LFC.



(a)



(b)

Fig. 4. (a) Power deviation in generator with bilateral contract given random packet delays. (b) System frequency response with bilateral contracts given random packet delays.

The packet delays are the sum of delays on each subnet link traversed by the packet, with each link delay consisting of processing delay, queueing delay, transmission delay and propagation delay [6]. The effects of retransmission are neglected since they are rare for most links. The models give a good approximation for LFC communication using UDP for sending the control signals. The focus of this paper is mostly on two scenarios, namely, a dedicated star topology for the traditional AGC model and a distributed model based on a dedicated network configuration. The latter also applies to the nondedicated distributed structure. A good background in queueing models can be found in [7]–[11].

A. Case 3.1: Star Topology

The traditional AGC model consists of dedicated links all emanating from the control center, now ISO, which processes the signal and sends them to the respective generators. The signal packets are all of the same length and are sent out say, every 4 seconds. This conforms to a D/G/1 model, with the arrival packet distribution deterministic and the service times being some general distribution due to the variety at the outgoing links. The notation conforms to Kendall's notation [6]. Note, further simplification to a D/D/1 model can be assumed if the outgoing links are assumed identical, in which case, the result is a trivial form of the G/G/1 model. Queueing theory models are all founded on three main entities,

- the arrival process denoted by the first alphabet, which in this case refers to a general arrival distribution (e.g., D deterministic, G general),
- the service process denoted by the second alphabet, and
- the number of servers or queues denoted by the last number.

Assume that the interarrival times at the queue are independent of each other and that the service times are independent identically distributed (i.i.d), meaning that they are independent of interarrival times and each packet service is mutually independent of each other. Let one further assume that all entities (such as waiting time, number of packets in the system, etc.) reach a steady state value. These assumptions should hold as the signals sent to each generators are independent and the outgoing links are independent of the neighboring channels. Note that in the case when the arriving signal packets are indeed dependent on each other (in case of concurrent signals being sent to the same site but differing generators), modeling can be done using either batch arrival models [8] or ON-OFF [2] processes. This is beyond the scope of this paper.

For the AGC star model, it can be shown that the average waiting time in queue asymptotically approaches

$$W \leq \frac{\lambda(\sigma_a^2+\sigma_s^2)}{2(1-\rho)} \tag{1}$$

where σ_a^2 is the variance of the interarrival time, σ_s^2 is the variance of the service time, λ is the average interarrival time, $1/\mu$, is the average service time, $\rho = \rho/\mu$ the utilization factor.If the $k+1$ packet arrives when the queue is empty, then its waiting time is 0 otherwise it is equal to the difference of the time taken to process the kth packet and the time for the $k+1$ packet to arrive. Thus

$$W_{k+1} = \max\{0, W_k + X_k - \tau_k\} \tag{2}$$

where W_k is the waiting time of the kth customer, X_k is the service time of the kth customer, and τ_k is the interarrival time between packets k and $k+1$. The idle time of the server is 0 if the $k+1$ packet arrives during the processing of the kth packet, otherwise it is equal to the difference between its arrival and the processing time for the kth packet. Now, let so that

$$I_k = (W_k + V_k)^- \tag{3}$$

I_k is referred to as the idle period length between packets k and $k+1$. For simplicity, generically denote $Z^+ = \max\{0, Z\}$, $Z^- = -\min\{0, Z\}$, $\overline{Z} = E\{Z\}$ and $\sigma_z^2 = E\{Z^2 - \overline{Z}^2\}$ [6]. It can be easily seen that $\overline{Z} = \overline{Z}^+ - \overline{Z}^-$ and so $\sigma_z^2 = \sigma_{z^+}^2 + \sigma_{z^-}^2 + 2\overline{Y}^+\overline{Y}^-$. Thus

$$\sigma_{I_k}^2 = \sigma_{\overline{w}_{k+1}}^2 + \sigma_{\overline{V}_k}^2 - 2\overline{W}_{k+1}\overline{V}_k \tag{4}$$

with W_k and V_k are independent. Additionally, arrival time and service time are independent, so

$$\sigma_{\overline{V_k}}^2 = \sigma_a^2 + \sigma_s^2. \tag{5}$$

Taking the limit as $k \to \infty$, and assuming that steady state values exist (i.e., assumption two), $\overline{W_k} \to W$, $\overline{I_k} \to I$. The average idle time between two successive arrivals is equal to the fraction of the time the system is idle, multiplied by the average interarrival time, that is, $I = (1-\rho)/\lambda$. Substituting this in (4) and combining with (5)

$$W = \frac{\lambda(\sigma_a^2+\sigma_s^2)}{2(1-\rho)} - \frac{\lambda\sigma_I^2}{2(1-\rho)}. \tag{6}$$

Note as the system gets heavily loaded $\sigma_I^2 \to 0$ so

$$W \leq \frac{\lambda(\sigma_a^2+\sigma_s^2)}{2(1-\rho)}. \tag{7}$$

For any packet, the delay is the summation of the average waiting time and the average service time. Thus

$$T \leq \frac{\lambda(\sigma_a^2+\sigma_s^2)}{2(1-\rho)} + \frac{1}{\mu} \tag{8}$$

where T is the steady state delay. For deterministic arrival, $\sigma_a^2 = 1/\lambda^2$ so

$$T \leq \frac{\frac{1}{\lambda}+\lambda\sigma_s^2}{2(1-\rho)} + \frac{1}{\mu}. \tag{9}$$

Note Kingman's heavy traffic approximation [6] that states under heavy traffic the steady state waiting time distribution can be approximated by an exponential distribution with mean $\lambda(\sigma_a^2+\sigma_s^2)/2(1-\rho)$.

B. Case 3.2: Star Topology- Server Outages or Denial of Service

This case is akin to the modeling of a server with non- exhaustive vacations and FIFO service, which means that the vacation time can occur at any point whether it is busy or not. The modeling of this is made simple by the condition of independence of the arrival processes from the number of packets already present in the system, which is assumed to hold for our system, since typically the arrival of the packets do not depend on the queue size. The situation where this does not hold, is in the case of smart routing, where, depending on the load on each node, the routing protocol redirects the packet to another route. Under the independence assumption, the distribution of the number of messages in the system exhibits a stochastic decomposition property at message departure time [10]. This effectively means that the number of packets present before the start of a vacation, the number that arrive during a vacation, and the number at any arbitrary time are independent of each other. Thus, the results of each can be calculated independently and added to give the total delay. The decomposition is

$$\pi(z) = H_{-}(z)\alpha_{-}(z)\pi_{G/G/1}(z). \tag{10}$$

That is, with π_k the probability of k messages in the system immediately after a departure time, then let $\pi(z)$ be the generating function for π_k, $H_{-}(z)$ be the probability generating function (PGF) for the number of messages in the system at the beginning of the vacation period, $\alpha_{-}(z)$ be the PGF for the number of messages that arrive before an arbitrary message during a vacation period and $\pi_{G/G/1}(z)$ the $\pi(z)$ for a G/G/1 system without vacations.

The expected message delay is

$$E[T] = \frac{E[L^{-}]}{\lambda} + \frac{\alpha^{(2)}(1)}{2\lambda E[\alpha]} + \frac{\lambda b^{(2)}}{2(1-\rho)} + b \tag{11}$$

where L^{-} is the number of messages at the beginning of a vacation period, α is the number of messages that arrive during a vacation period, λ is the message arrival rate, b is the mean packet service time and again ρ is the utilization factor.

C. *Case 3.3: Distributed Dedicated Network*

A dedicated network offers a strong fault tolerance guarantee as well as low latency and variation on packet delivery. The analysis of delay is nontrivial as tandem queues are correlated and hence, the assumptions of interdependence of service times and arrival times break down. Under certain assumptions, some simplifications can be made to decompose the individual queues. Several empirical studies performed on large packet monitors have determined the arrival process for certain internet traffic to be statistically self similar with a log normal distribution with a heavy tail [12]. The data signal for load following should not conform to that distribution but could be more simply modeled as a G/G/1 queue model at each intermediate queue. This is because the signals are sent every few seconds all at once and hence their arrivals at the source queue are deterministic. This type of data has been shown to follow a more Poisson distribution characteristic [12].

From Jackson's theorem (extended to G/G/1 queue) [8], [9], [11] and the Kleinrock independence approximation [11], the system of tandem queues can be effectively decomposed to be an independent set of G/G/1 queues. Hence, the delays can be approximated well as the summation of the delays the packet encounters during its route with each node being modeled as a G/G/1 queue. This is dependent on the route the data packet eventually takes, which cannot be determined before hand. Still with signals sent every few seconds, the path traversed would eventually resort to the shortest path to the destination and the delays determined for this particular path would hold. For the scenario where bandwidth reservation is implemented, the path is effectively a virtual circuit network and thus, the delays can be calculated effectively.

Finally, for the situation where one needs to determine the effect of server outage or denial of service at any one node, the distributed nature of the network resorts to redirecting the signal along a different path, and effectively, the delay variation should not be large. In the case when such outages occur in a large part of the network or at the source, then the delays can be approximated assuming that the vacation time lends itself to decomposition.

IV. Discussion

In the near future, the power system's communication system will inevitably face contingencies similar to those presented in the Section II. As shown, a problem in the communication system can compromise the system integrity. Most often the anomaly slows down the system response, but in the worst case, it can lead the system toward instability or other unacceptable behavior. Characterizations of the communication signal delays are important to model and can allow fault tolerant controls to be developed. This will become even more critical for fast time dependent applications, such as, stability controls. Even though the load following signals are typically not considered critical a relatively simple denial of service attack perpetrated by any individual on this service can lead to severe problems. A reliant and robust communication network with low delays and small delay variation is certainly needed. Some emergent technologies, such as, bandwidth reservation and cognizant routing techniques, could be used to provide good quality of service for the critical data packets.

Delay bounds also help in providing guidelines for intrusion detection and overall failure detection in the LFC system. The security of signal data is a necessity in an open communication scenario and should be strictly adhered to and implemented by all players in the load following ancillary service. Strong encryption is probably not a requirement as the security expiry time is on the order of a few seconds. A possible exception arise if there are situations where a competitor monitored a sequence of data, if not encrypted, that then led to some competitive advantage. This is not viewed as likely. Some form of data authentication should also be implemented to determine the source of the data before making the indicated change in generation.

These communication network studies are germane to the ISO and even more so to third party participants as they are the most susceptible to delays in the communication system. Such models could help contracts guarantee some level of quality to their customers for a given load following service.

V. Conclusion

This paper demonstrates the need for including the effects of communication delays in LFC for a deregulated market. Section II shows the effect of constant delays as well as random delays introduced in the dissemination of the LFC signals to the participating generators. The simulations of Section II demonstrate failures due to these delays. Section III builds simple models in an effort to provide approximate values for the delays to characterize a given communication network. Values for the equations shown are dependent on current network traffic characteristics, traffic load as well as the particular topology of the network. Data for this can be obtained by sampling the network data traffic at pertinent communication nodes over short as well as long periods of time. These would then provide the input to the LFC models.

The simple models derived in Section III provide guidance for subsequent control design and were used in the power system simulations here. Still, practical characterization of a specific large distributed communication network can only be obtained through simulations involving all the communication network parameters. This can be performed using widely available network simulators (such as NS-2 or Opnet), where the inclusion of very detailed models, e.g., protocol delay characteristics, is also possible.

Acknowledgment

The authors would like to thank Dr. Shivalingam and Dr. M. Vaziri for their suggestions and helpful information.

References

[1] NERC Operating Manual, Policy 10. North American Electric Reliability Council. [Online]. Available: http://www.nerc.com

[2] R. J. Adler, R. Feldman, and M. S. Taqqu, *A Practical Guide to Heavy Tails: Statistical Techniques and Applications*. New York: Springer Verlag, 1998.

[3] E. Nobile, A. Bose, and K. Tomsovic, "Feasibility of a bilateral market for load following," *IEEE Trans. Power Syst.*, vol. 16, pp. 782–787, Nov. 2001.

[4] A. J. Wood and B. F. Wollenberg, *Power Generation Operation and Control*. New York: Wiley, 1984.

[5] S. Bhowmik, "Effect of Communication Network Infrastructure on Load Frequency Control," M.S., Washington State Univ., Pullman, WA, 2000.

[6] Bertsekas and R. Gallager, *Data Networks*. Englewood Cliffs, NJ: Prentice-Hall, 1992.

[7] A. Papoulis, *Probability Random Variables and Stochastic Processes*. New York: McGraw-Hill, 1991.

[8] E. Gelenbe and G. Pujolle, *Introduction to Queueing Networks*. New York: Wiley, 1998.

[9] J. Mehdi, *Stochastic Models in Queueing Theory*. New York: Academic, 1991.

[10] B. T. Doshi, "Queueing systems with vacations -a survey," *Queueing Systems*, vol. V1, pp. 29–66, 1986.

[11] L. Kleinrock, *Queueing Systems*. New York: Wiley, 1975.

[12] W. E. Leland, M. S. Taqqu, W. Willinger, and D. V. Wilson, "On the self similar nature of ethernet traffic (Extended Version)," *IEEE/ACM Trans. Networking*, vol. 2, pp. 1–14, 1994.

Sudipto Bhowmik received the B.E. degree from Rashtriya Vidyalaya College of Engineering with distinction in 1996 and the M.S. degree from Washington State University, Pullman, in 2000.

Currently, he is with Network Elements Inc., Beaverton, OR. He was with Asea Brown Boveri, Bangalore, India, from 1996 to 1998.

Kevin Tomsovic (SM'00) received the B.S. degree in electrical engineering from Michigan Tech. University, Houghton, in 1982, and the M.S. and Ph.D. degrees in electrical engineering from the University of Washington, Seattle, in 1984 and 1987, respectively.

Currently, he is an Associate Professor at Washington State University, Pullman. Visiting university positions have included Boston University, Boston, MA; National Cheng Kung University, Tainan, Taiwan, R.O.C.; National Sun Yat-Sen University, Kaohsiung, Taiwan, R.O.C., and the Royal Institute of Technology, Stockholm, Sweden. He held the Advanced Technology for Electrical Energy Chair at Kumamoto University, Kumamoto City, Japan, from 1999 to 2000.

Anjan Bose (M'68–SM'77–F'89) received the B.tech (Hons.) degree from the Indian Institute of Technology, Kharagpur, in 1967, the M.S. degree from the University of California, Berkeley, in 1968, and the Ph.D. degree from Iowa State University, Ames, in 1974.

Currently, he is the Distinguished Professor in Power Engineering and Dean of the College of Engineering and Architecture at Washington State University, Pullman. He has was with Consolidated Edison Co., New York, from 1968 to 1970; the IBM Scientific Center, Palo Alto, CA, from 1974 to 1975; Clarkson University, Potsdam, NY, from 1975 to 1976; Control Data Corporation, Minneapolis, MN, from 1976 to 1981; and Arizona State University, Tempe, from 1981 to 1993.