Modeling and Characterization of Traffic in Public Safety Wireless Networks

Božidar Vujičić, Nikola Cackov, Svetlana Vujičić, and Ljiljana Trajković*
Simon Fraser University
Vancouver, BC, Canada
{bvujicic, ncackov, svujicic, ljilja}@cs.sfu.ca

Keywords: Trunked radio systems, wireless networks, traffic analysis, long-range dependence, wavelets.

Abstract
In this paper, we present statistical analysis of traffic in a deployed circuit-switched, trunked cellular wireless network used by public safety agencies. Prior analysis considered traffic from three busy hours during two days in 2001. Traffic data in this study span various time periods in 2001, 2002, and 2003. We examine the statistical distribution and autocorrelation function of call inter-arrival and call holding times during several busy hours. We find that call inter-arrival times are long-range dependent and may be modeled by both Weibull and gamma distributions. Call holding times follow the lognormal distribution and are uncorrelated. These findings indicate that traditional Erlang models for voice traffic may not be suitable for evaluating the performance of trunked radio networks.

1. INTRODUCTION

Analysis of traffic in deployed communication networks is important for determining their operational status. Traffic modeling is necessary for network provisioning, predicting utilization of network resources, and for planning network developments. Reliable operation is particularly important for networks used by public safety agencies, such as police, fire department, and ambulance [1]–[3].

E-Comm, Emergency Communications for South West British Columbia Inc. [4], provides voice and data transmission services via a Public Safety Wireless Network to several public safety agencies. The E-Comm network is a circuit-switched macro-cellular network [3]. The cells cover various areas of the Greater Vancouver Regional District. Each of the eleven cells has a number of available radio channels, which determines its capacity. Access to the channels is based on the concepts of trunking.

Previous results on performance evaluation of cellular trunked radio systems have been obtained by creating mathematical models based on queuing theory [5] and statistical analysis of collected traffic [6]–[8]. The channel holding time in the analyzed micro-cellular network, where handoffs are frequent, follows the lognormal distribution [7]. The Public Access Mobile Radio (PAMR) system [8] is a macro-cellular network similar to the E-Comm network. It employs a push-to-talk mechanism for network access and is used by several groups working on transport and distribution. The PAMR system uses message trunking, where the entire conversation is treated as one call. The analysis of message and transmission durations revealed that message length (channel holding time) and transmission length may be modeled by Erlang-*jk* and lognormal (or a mixture of two lognormal) distributions, respectively [8].

Performance of the E-Comm network has been evaluated by using the OPNET network simulator [1]. Simulations of the network utilization during two sample weeks also addressed the increase in the network traffic volume and network congestion. A customized simulation tool and an initial statistical analysis of traffic from 2001, 2002, and 2003 have been also reported [3]. Furthermore, clustering of the network users based on their activity and a Seasonal ARIMA model for predicting traffic from each cluster were proposed in [9]. A statistical analysis of the E-Comm network traffic from three busy hours in 2001 was presented in [2]. In this paper, we extend this statistical analysis by considering traffic traces from several busy hours collected over various periods during three years and by testing the goodness-of-fit of a larger number of probability distributions.

Voice traffic in circuit-switched networks has been modeled by the Erlang B and C models [10]. The models assume independent and exponentially distributed call holding and call inter-arrival times. They have proved appropriate for modeling telephone traffic. Trunked radio systems possess characteristics that distinguish them from the telephone networks, such as trunking-based network access and one-to-many type of conversation. Therefore, the Erlang models may not capture the statistical characteristics of the traffic in trunked radio systems. Our analysis of traffic from the E-Comm network shows that neither call holding nor call inter-arrival times are exponentially distributed. Furthermore, call inter-arrival times exhibit long-range dependence.

In Section 2, we describe the structure and operating of the E-Comm network. Statistical concepts and tools

* This research was supported by the NSERC Grant 216844–03 and Canada Foundation for Innovation.

employed for the traffic analysis are presented in Section 3. Traffic data, analysis results, and comparison of the parameters of traffic from various years are shown in Section 4. We conclude with Section 5.

2. DESCRIPTION OF THE E-COMM NETWORK

E-Comm public safety wireless network is based on the Enhanced Digital Access Communication System (EDACS) [11]. The EDACS architecture is shown in Fig. 1. Its main elements are the central system controller (network switch), radio repeaters, fixed user sites (dispatch consoles), mobile users, and a management console. EDACS is connected to the public switched telephone network (PSTN) via a public branch exchange (PBX) gateway and to packet networks via the data gateway. Network activity events are recorded in a central database.

The wireless section of the E-Comm network has a cellular architecture. Each cell is covered by one or more radio repeaters, depending on the cell's area. The transmission method between the repeaters and the mobile users is simulcast, which implies that all repeaters belonging to one cell use an identical set of frequencies. The number of frequencies in a cell determines the number of available radio channels and the number of simultaneous radio transmissions (calls). Cell's capacity is determined based on the expected traffic volume.

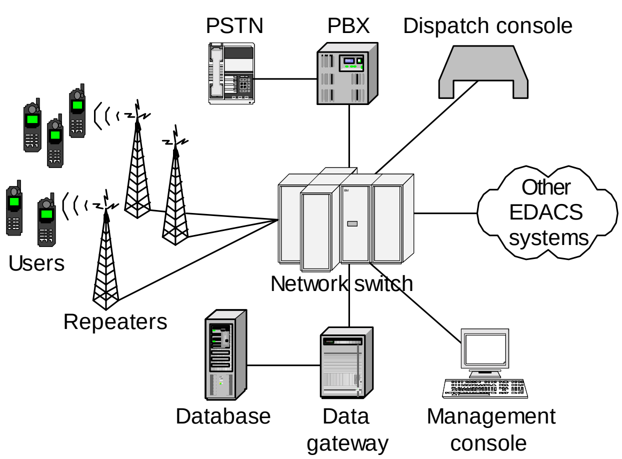

Fig. 1. EDACS architecture [3].

The management of frequencies (radio channels) in the E-Comm network is based on trunking. It implies sharing all frequencies in a cell among all agencies instead of dedicating subsets of frequencies to individual agencies. This approach results in better utilization of radio resources and minimizes the number of radio channels necessary for matching certain Grade of Service (GoS) requirements. The E-Comm network utilizes transmission trunking (each radio transmission is treated as a separate call) rather than message trunking. With transmission trunking, when a user presses the push-to-talk (PTT) button on the mobile radio transceiver, a call is established (a channel is dedicated to the transmission). The call lasts as long as the user holds the PTT button. The call ends and the radio channel is released when the PTT button is released.

Users of the E-Comm network are organized in talk groups that belong to agencies such as police, fire department, and ambulance. A user may be a member of more than one talk group and may switch between talk groups dynamically. The most common type of a call in the network is the group call. This implies a one-to-many type of conversation: a user talks to all other members in his/her talk group. Depending on the locations of the members of the talk group, a call may require one or more channels. If all members of the talk group are within one cell, the call is established using one channel. If there are members of the talk group in several cells, a channel in each destination cell must be dedicated to the call. When no channel is available in at least one destination cell, the call is queued. If the call cannot be established within a certain time period, it is dropped.

Each cell has a distinct pool of frequencies. The number of frequencies determines its capacity. Cells, with their predefined and limited capacities, are main network bottlenecks. Queued and dropped calls occur due to the insufficient number of radio channels in the cells. Therefore, analyzing and modeling call traffic from each cell individually is important to determine current and predict future network performance. Each cell covers a relatively large area (entire municipality) and the calls are relatively short (average 3.8 s [3]). This implies rare occurrence of call handover.

3. STATISTICAL CONCEPTS AND ANALYSIS TOOLS

Statistical processes possess two important characteristics: probability distribution and autocorrelation. The probability distribution characterizes the probability that the outcomes of the process (random variables) are within a given range of values. It is expressed through probability density and cumulative distribution functions. Probability density functions show the probability of occurrence of a certain value or range of values. Cumulative distribution functions express the probability that the variable will not exceed specific values. Autocorrelation function measures the dependence between two outcomes of the process. In general, it is a function of the time instances of the two outcomes. If the process is wide-sense stationary, its autocorrelation function depends only on the difference (lag) between the time instances of the outcomes.

The traffic in the E-Comm network is characterized by two processes: call arrival and call holding processes. Outcomes of the call arrival and call holding processes are the sequences of call inter-arrival and call holding times, respectively. Investigating the statistical properties (probability distributions and the autocorrelations) of these processes is important for deriving an appropriate traffic model and employing the model for determining network performance. Choosing the statistical distribution that best fits the data is performed by comparing the distribution of the data with several known distributions and employing the Kolmogorov-Smirnov (K-S) goodness-of-fit test. The autocorrelation of the data is examined by plotting the autocorrelation functions and testing whether the data exhibit long-range dependence.

3.1 Long-Range Dependence (LRD)

For mathematical simplicity, it is often assumed that a process is wide-sense stationary and uncorrelated, or that its autocorrelation function is zero for non-zero lags. This assumption does not hold for all processes. Often, there is a certain correlation structure that cannot be neglected. A class of processes with non-negligible autocorrelations is the family of long-range dependent (second-order self-similar) processes [12], [13]. Long-range dependence is defined as a non-summability of the autocorrelation function $r(k)$ of a wide-sense stationary process $X(n)$, $n = 1, 2, 3, \ldots$. The autocorrelation function $r(k)$ of an LRD process is modeled as a hyperbolically decaying function

$$r(k) = c_r k^{-(2-2H)}, \ \ k \to \infty, \tag{1}$$

where c_r is a positive constant and H ($0.5 \leq H < 1$) is the Hurst parameter. The power spectral density (PSD) $f(v)$ of $X(n)$ satisfies

$$f(\nu) = c_f |\nu|^{-\alpha}, \ \ |\nu| \to 0, \tag{2}$$

where c_f is a positive constant and α is the scaling exponent [14]. For LRD processes $0 < \alpha < 1$ and the relationship between H and α is linear:

$$H = 0.5(1+\alpha). \tag{3}$$

The Hurst parameter H measures the degree of LRD of a process. Values of $H \approx 1$ imply strong LRD (strong correlations between outcomes of the process that are far apart). For uncorrelated processes $H = 0.5$.

3.2 Wavelets and Wavelet-Based Estimator of *H*

The discrete wavelet transform (DWT) of a signal $X(t)$ is given by the inner product

$$d(j,k) = \int_{-\infty}^{\infty} X(t)\psi_{j,k}(t)dt, \tag{4}$$

where $d(j, k)$ is the wavelet coefficient at octave j and time k and

$$\psi_{j,k}(t) = 2^{-j/2}\psi\left(2^{-t}t - k\right), \quad j \in Z^{+}, k \in Z \tag{5}$$

is the basis function called a wavelet. It is obtained by scaling (by a factor of 2^{-j}) and translating (by k time units) of an adequately chosen mother wavelet ψ [14]. The mother wavelet possesses two important properties. It is an oscillating function (its mean value is zero). Furthermore, most of its energy is concentrated within limited time interval and limited frequency band. The signal $X(t)$ is represented as a weighted sum of wavelets:

$$X(t) = \sum_{j=0}^{\infty}\sum_{k=-\infty}^{\infty} d(j,k)\psi_{j,k}(t). \tag{6}$$

DWT captures signals over various time scales. Wavelets' scale invariance makes DWT suitable for analyzing properties that are present across a range of time scales, such as LRD. Furthermore, the existence of low computational cost algorithms for implementing DWT makes DWT a popular tool for signal analysis.

The wavelet-based Hurst parameter estimator is based on the shape of the PSD function (2) of the LRD signal $X(t)$. It has been shown [14] that the power-law behavior of the PSD implies the following relationship between the variance of the wavelet coefficients and the octave j:

$$\mathrm{E}\left\{d(j,k)^2\right\} = 2^{j\alpha} c_f C(a,\psi), \tag{7}$$

where the average is calculated for various k, α is the scaling exponent, and $C(\alpha,\psi)$ depends on the mother wavelet, but does not depend on j. When the mother wavelet is suitably chosen [14], $\mathrm{E}\{d(j,k)^2\}$ is a sample mean of $d(j,k)^2$ calculated over all k's:

$$\mathrm{E}\left\{d(j,k)^2\right\} = \frac{1}{n_j}\sum_{k=1}^{n_j} d(j,k)^2, \tag{8}$$

where n_j is the number of coefficients available at octave j. The plot of $\log_2\mathrm{E}\{d(j,k)^2\}$ vs. j is called a *logscale diagram*. Linear relation with a slope α ($0 < \alpha < 1$) between $\log_2\mathrm{E}\{d(j,k)^2\}$ and j for a range of octaves, including the coarsest, indicates presence of LRD. Therefore, α is obtained by performing linear regression of $\log_2\mathrm{E}\{d(j,k)^2\}$ on j over a range of octaves. The Hurst parameter is obtained from (3).

We employed the publicly available MATLAB code [15] to estimate H. In the analysis, we used the wavelet Daubechies3.

3.3 Test for Time Constancy of the Scaling Exponent *α*

LRD processes are, by definition, wide-sense stationary. However, they possess certain characteristics that make them seem non-stationary. LRD processes exhibit high variability [16] and there are relatively long on and off periods. An important issue is how to distinguish between wide-sense stationary processes with LRD and inherently non-stationary processes.

An approach to determine whether a process is LRD or non-stationary is to test if the scaling exponent α is constant over the examined time series [16]. Time constancy of α is also important because the wavelet-based estimator may produce unreliable estimates when applied to time series with variable α. The test for time constancy of α [15] divides the time-series into m blocks of equal lengths and estimates α for each block. The estimates are compared and a decision is made whether or not α can be considered constant over the duration of the entire time series.

3.4 Kolmogorov-Smirnov Test

Kolmogorov-Smirnov (K-S) goodness-of-fit test is employed to determine the best fit among several distributions [17]. The null hypothesis H_0 implies that data samples follow a given distribution. The alternative hypothesis H_1 states the opposite. The purpose of the test is to check whether to accept or reject the null hypothesis H_0 and to quantify the decision. The approach of the K-S test is to examine whether the empirical distribution of a set of observations (empirical cumulative distribution function) is consistent with a random sample from an assumed theoretical distribution. The empirical cumulative distribution function E_N is defined as a step function (with step size $1/N$) of N ordered data points $Y_1, Y_2, ..., Y_N$:

$$E_N = \frac{n(i)}{N}, \tag{10}$$

where $n(i)$ is the number of data samples with values smaller than Y_i.

The decision whether or not to accept the null hypothesis H_0 is based on the value of the test statistic k, defined as the maximum difference over all data points:

$$k = \max_{1 \le i \le N} \left| F(Y_i) - \frac{i}{N} \right|, \tag{11}$$

where F is the cumulative distribution function of the assumed distribution. It means that for each data point, the K-S test compares the proportion of values less than that data point with the number of values predicted by the assumed distribution. The null hypothesis is accepted if the value of the test statistic is lower than the critical value. Three additional parameters play an important role in analyzing the test results. The significance level α (default value equals 0.05) determines that the null hypothesis is rejected α percent of the times when it is in fact true. It defines the sensitivity of the test. Smaller values of α imply larger tolerance (larger critical values). The second parameter *tail* specifies whether the K-S performs a two sided test (default) or alternative tests from one or other side. The third parameter is the observed p-value that reports the probability level on which the difference between distributions (test statistics) becomes significant. If $p \le \alpha$, the test rejects the null hypothesis. Otherwise, the null hypothesis is accepted. Parameters α and *tail* are input parameters and the p-level is one of the test results. If the test returns a non-number for the critical value, then the decision to accept or reject the null hypothesis is based only on the p-value [17].

A difficulty in applying goodness-of-fit tests is that results are dependent on the sample size [17]. It is not uncommon for the test to reject the null hypothesis when large datasets are tested. A solution is to perform the test on randomly chosen subsets of data.

4. ANALYSIS OF DATA

We analyze traces of call holding and call inter-arrival times. For each trace, we determine the best fitting distribution using the K-S test and examine the autocorrelation function. We also test the traces for long-range dependence by performing wavelet-based estimation of the Hurst parameter and by testing the time constancy of the scaling exponent α. Figures included in the paper illustrate graphical results obtained from the 2003 busiest hour (between 22:00 and 23:00 on 26 March). The results from the other hourly traces are similar.

4.1 Traffic Data

Traffic data from E-Comm consist of records of network events, such as established, queued, and dropped calls. Each established call is identified by its timestamp, duration, caller, callee, and destination cell(s). From the traffic data, we create traces of call holding times (call durations) and call inter-arrival times (differences between two successive timestamps). We analyze traces from the cell covering Vancouver because it is the busiest cell and handles the majority of the calls. It also has the largest number of available radio channels and a sufficient capacity so that congestion and call queuing rarely occur.

Analyzed traffic traces (datasets) span various periods during three years: 2001, 2002, and 2003. The time span and the number of calls in each dataset are shown in Table 1. We determine the number of calls in every one-hour interval of each dataset in order to identify the busiest hours. Our analysis focuses on call holding and call inter-arrival times from the five busiest hours in each dataset. Analysis of busy hour traffic is typical for circuit-switched networks because they are designed to satisfy certain Grade of Service (GoS) requirements regarding the frequency of occurrence and duration of call queuing during periods of high utilization [18]. The number of calls during the busiest hours in each dataset is shown in Table 2.

Table 1. Time span and number of calls in the traffic traces.

Trace (dataset)	Time span	No. of calls
2001	November 1–2, 2001	110,348
2002	March 1–7, 2002	370,510
2003	March 24–30, 2003	387,340

Table 2. Five busiest hours in the traffic traces from 2001, 2002, and 2003 with the corresponding number of calls.

2001		2002		2003	
Day/hour	No.	Day/hour	No.	Day/hour	No.
02.11.2001 15:00–16:00	3,718	01.03.2002 04:00–05:00	4,436	26.03.2003 22:00–23:00	4,919
01.11.2001 00:00–01:00	3,707	01.03.2002 22:00–23:00	4,314	25.03.2003 23:00–24:00	4,249
02.11.2001 16:00–17:00	3,492	01.03.2002 23:00–24:00	4,179	26.03.2003 23:00–24:00	4,222
01.11.2001 19:00–20:00	3,312	01.03.2002 00:00–01:00	3,971	29.03.2003 02:00–03:00	4,150
02.11.2001 20:00–21:00	3,227	02.03.2002 00:00–01:00	3,939	29.03.2003 01:00–02:00	4,097

Fig. 2 shows a time series of the call traffic between 22:18 and 22:19 on March 26, 2003 (one minute of the busiest hour in the 2003 dataset). The horizontal axis shows the timestamps of the call. The vertical axis shows the call holding times. Call inter-arrival times are observed as time intervals between two successive calls.

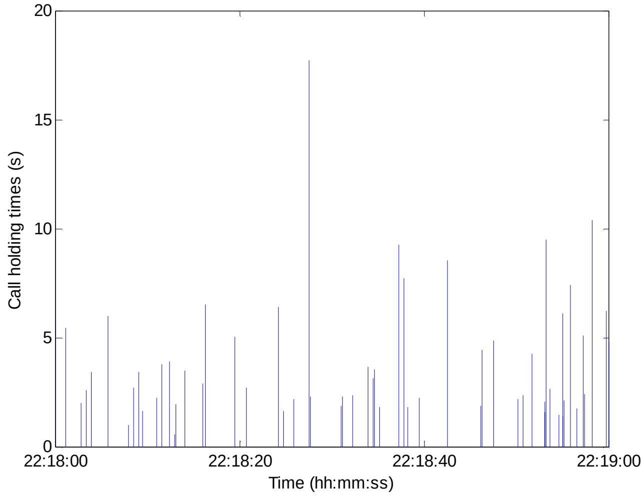

Fig. 2. Time series of the one-minute interval between 22:18 and 22:19 from the 2003 busiest hour traffic.

4.2 Call Inter-Arrival Times

We consider the following fourteen candidate distributions: exponential, Weibull, gamma, normal, lognormal, logistic, log-logistic, Nakagami, Rayleigh, Rician, *t*-location scale, Birnbaum-Saunders, extreme value, and inverse Gaussian. The parameters of the distributions are calculated by performing Maximum Likelihood Estimation using the hourly traces of call inter-arrival times. For each trace, we plot the distribution of the trace (histogram) and the probability density function for each distribution. Examples of the plots are shown in Fig. 3.

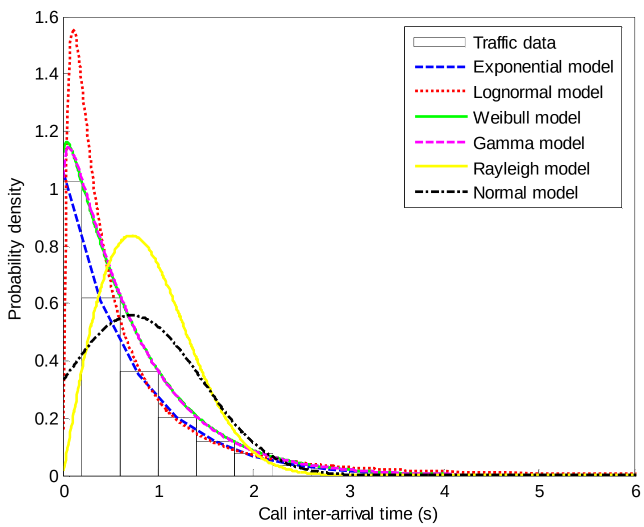

Fig. 3. Call inter-arrival times distributions.

It can be concluded that exponential, lognormal, Weibull, and gamma distributions fit the data better than the others. Quantitative results of the two-sided K-S goodness-of-fit test for the 2001, 2002, and 2003 datasets are shown in Tables 3, 4, and 5, respectively. The test is performed on the entire traces. The significance level α is 0.1. The null hypothesis (distribution fits the data) is rejected if the test statistics k is greater than the critical value. In that case, the test returns h=1. Otherwise, the test returns h=0, which implies that the distribution fits the data

Both Weibull and gamma distributions fit the call inter-arrival times in E-Comm network, as indicated in Tables 3, 4, and 5. The test accepts the null hypotheses with a significance level of 0.1 for most hourly traces. Weibull distribution fails the test for two traces in 2001. Gamma fails the test for three traces in 2001. We repeat the K-S test for those traces and the corresponding distributions for various values of α. The distributions may be accepted for smaller significance levels (between 0.01 and 0.1). In contrast, test results show that the lognormal distribution cannot be accepted as a suitable model because it did not pass the test for any trace and for any α. The exponential distribution also fails the test for the majority of traces. Nevertheless, its p-values are significantly higher than those of the lognormal distribution, which indicates a better fit.

Fig. 4 shows the cumulative distribution function of the call inter-arrival times from the 2003 busiest hour and the exponential, Weibull, and gamma distributions. The four curves almost overlap, indicating a good fit.

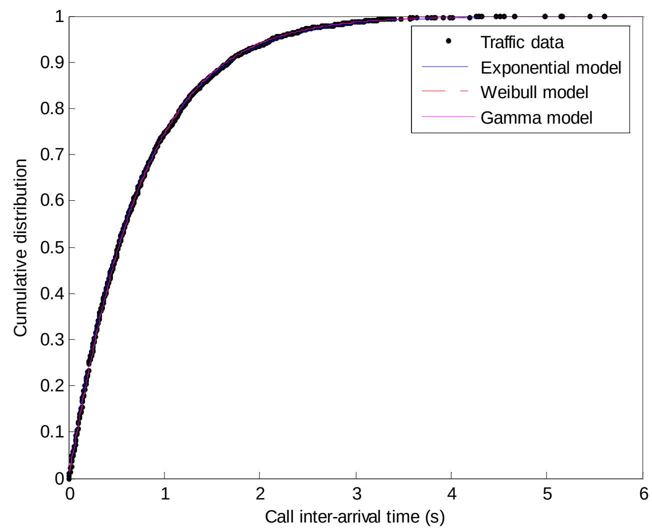

Fig. 4. Cumulative distribution function of the call inter-arrival times and comparison with exponential, Weibull, lognormal, and gamma distributions.

The autocorrelation function of the call inter-arrival times from the 2003 busiest hour is shown in Fig. 5. The horizontal lines show the 95% (dashed line) and 99% (dash-dotted line) confidence intervals, respectively. The majority of the autocorrelation coefficients for lags smaller than 60 are outside the confidence intervals. This indicates presence of non-negligible correlation among call inter-arrival times.

We test the traces of call inter-arrival times for long-range dependence by performing wavelet-based estimation of the Hurst parameter. An example of a logscale diagram is shown in Fig. 6. It exhibits a linear relationship between $\log_2 E\{d(j,k)^2\}$ and j with a positive slope in the range of

Table 3. K-S test results for the hourly traces of call inter-arrival times from the 2001 busy hours.

Distribution	Parameter	02.11.2001, 15:00–16:00	01.11.2001, 00:00–01:00	02.11.2001, 16:00–17:00	01.11.2001, 19:00–20:00	02.11.2001, 20:00–21:00
Exponential	h	1	1	0	1	1
	p	0.0384	1.45E-04	0.5416	0.0122	0.0135
	k	0.0247	0.0369	0.0131	0.0277	0.0259
Weibull	h	0	1	0	0	1
	p	0.3036	0.0409	0.4994	0.1574	0.0837
	k	0.0171	0.0236	0.0136	0.0195	0.0206
Gamma	h	0	1	0	1	1
	p	0.3833	0.0062	0.3916	0.0644	0.0953
	k	0.0159	0.0287	0.0148	0.0227	0.0202
Lognormal	h	1	1	1	1	1
	p	1.6520E-19	7.0722E-17	1.5936E-22	2.3731E-16	1.3828E-22
	k	0.0769	0.0713	0.0853	0.0743	0.0658

Table 4. K-S test results for the hourly traces of call inter-arrival times from the 2002 busy hours.

Distribution	Parameter	01.03.2002, 04:00–05:00	01.03.2002, 22:00–23:00	01.03.2002, 23:00–24:00	01.03.2002, 00:00–01:00	02.03.2002, 00:00–01:00
Exponential	h	1	1	1	1	1
	p	0.0105	6.58E-05	0.0089	0.0304	9.53E-04
	k	0.0243	0.0351	0.025	0.0229	0.0311
Weibull	h	0	0	0	0	0
	p	0.3865	0.6608	0.3384	0.8846	0.5918
	k	0.0136	0.0113	0.0143	0.0093	0.0123
Gamma	h	0	0	0	0	0
	p	0.2007	0.6704	0.5017	0.9135	0.4921
	k	0.0161	0.0112	0.0126	0.0089	0.0133
Lognormal	h	1	1	1	1	1
	p	3.0769E-19	2.4339E-16	9.9945E-16	1.7524E-18	4.0452E-19
	k	0.0698	0.0651	0.0648	0.0723	0.0739

Table 5. K-S test results for the hourly traces of call inter-arrival times from the 2003 busy hours.

Distribution	Parameter	26.03.2003, 22:00–23:00	25.03.2003, 23:00–24:00	26.03.2003, 23:00–24:00	29.03.2003, 02:00–3:00	29.03.2003, 01:00–02:00
Exponential	h	1	1	0	1	1
	p	0.0027	0.0469	0.4049	0.0316	0.1101
	k	0.0283	0.0214	0.0137	0.0205	0.0185
Weibull	h	0	0	0	0	0
	p	0.4885	0.4662	0.2065	0.286	0.2337
	k	0.013	0.0133	0.0164	0.014	0.0159
Gamma	h	0	0	0	0	0
	p	0.3956	0.3458	0.127	0.145	0.1672
	k	0.0139	0.0146	0.0181	0.0163	0.0171
Lognormal	h	1	1	1	1	1
	p	1.0147E-20	4.717E-15	2.97E-16	3.2665E-23	4.8505E-21
	k	0.0689	0.0629	0.0657	0.0795	0.0761

octaves [4–9], which indicates LRD. Estimates of H for all hourly traces are shown in Table 6. We also test the time constancy of the scaling exponent α by dividing each trace into m sub-traces, $m \in \{3, 4, 5, 6, 7, 8, 10\}$. All traces pass the test for more than 50% of m's, which indicates that α can be considered constant across the traces and the estimates of H reported in Table 6 are reliable. For all traces $H \in (0.5, 1)$, indicating that call inter-arrival times exhibit long-range dependence.

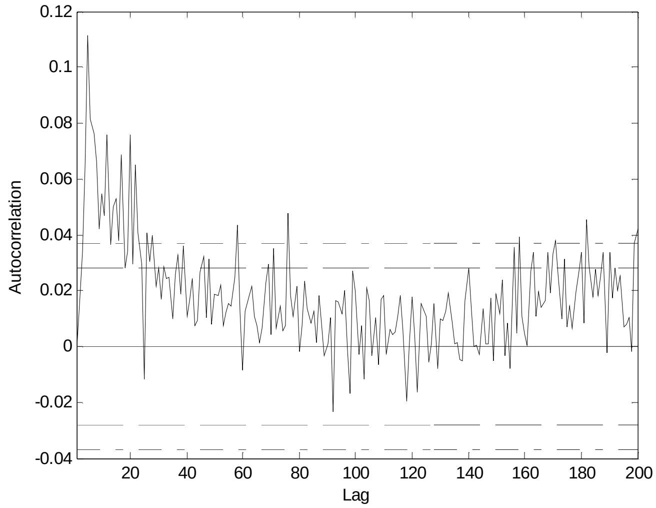

Fig. 5. Call inter-arrival times autocorrelation plot (up to lag 200) with 95% and 99% confidence intervals.

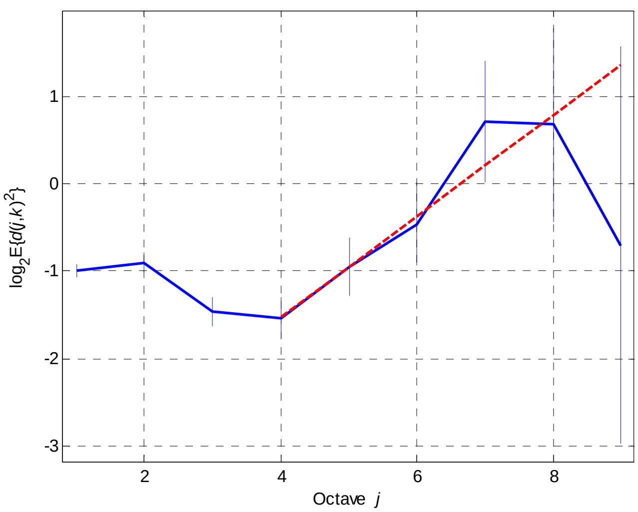

Fig. 6. Logscale diagram of the call inter-arrival times. Dashed line is the linear regression line with slope α. Vertical lines represent the 95% confidence intervals around the estimates of $\log_2 E\{d(j,k)^2\}$.

Table 6. Estimates of H for the hourly traces of call inter-arrival times.

2001		2002		2003	
Day/hour	H	Day/hour	H	Day/hour	H
02.11.2001 15:00–16:00	0.907	01.03.2002 04:00–05:00	0.679	26.03.2003 22:00–23:00	0.788
01.11.2001 00:00–01:00	0.802	01.03.2002 22:00–23:00	0.757	25.03.2003 23:00–24:00	0.832
02.11.2001 16:00–17:00	0.770	01.03.2002 23:00–24:00	0.780	26.03.2003 23:00–24:00	0.699
01.11.2001 19:00–20:00	0.774	01.03.2002 00:00–01:00	0.741	29.03.2003 02:00–03:00	0.696
02.11.2001 20:00–21:00	0.663	02.03.2002 00:00–01:00	0.747	29.03.2003 01:00–02:00	0.705

4.3 Call Holding Times

We compare the distribution of the call holding times with the distributions that were considered for the call inter-arrival times. The probability density function of the call holding times and of the several best fitting distributions are shown in Fig. 7. The K-S goodness-of-fit test is performed for the lognormal, exponential, Weibull, and gamma distributions. None of the distributions passes the test when the entire traces are tested with significance levels 0.1 and 0.01. Therefore, we perform the test on ten randomly chosen sub-traces of length 1,000 extracted from each trace, with a significance level α of 0.01. Only lognormal distribution passes the test for very few sub-traces. When sub-traces of length 500 are tested with the same significance level, the lognormal distribution exhibits the best fit. It passes the K-S test for almost all 500-sample sub-traces of all hourly traces. The test rejects the null hypothesis when those sub-traces are compared with the other three candidate distributions: exponential, Weibull, and gamma. Fig. 8 shows the cumulative distribution function of the call holding times from the 2003 busiest hour. It also shows the cumulative distribution function of the exponential, lognormal, Weibull, and gamma distributions. Lognormal distribution fits best the call holding times, as shown in Figs. 7 and 8.

The autocorrelation function of the call holding times from the busiest hour in the 2003 dataset is shown in Fig. 9. We observe that there are no significant correlations for non-zero lags because all but a few autocorrelation coefficients are within the 95% and 99% confidence intervals.

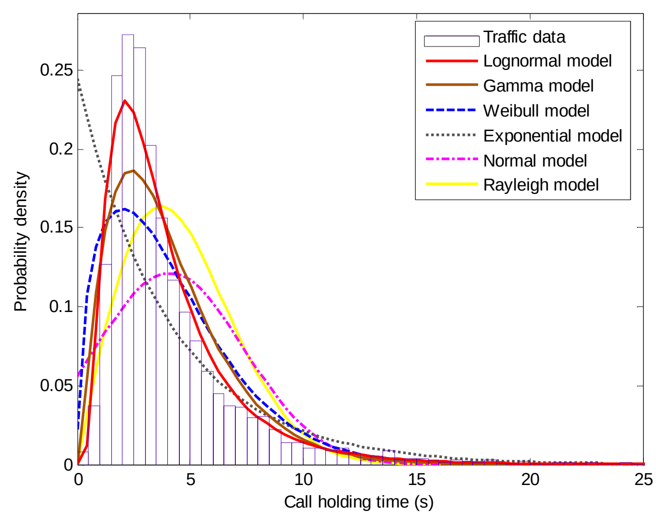

Fig. 7. Call holding times distributions.

We also investigate the long-range dependence in the call holding times. Fig. 10 shows an example of a logscale diagram for the call holding times. The linear region has a slope of approximately zero, which implies absence of LRD. The test for time constancy of the scaling exponent α indicates that all but one estimate of H can be considered reliable. The only unreliable estimate is from the busy hour between 23:00 and 24:00 on March 26, 2003.

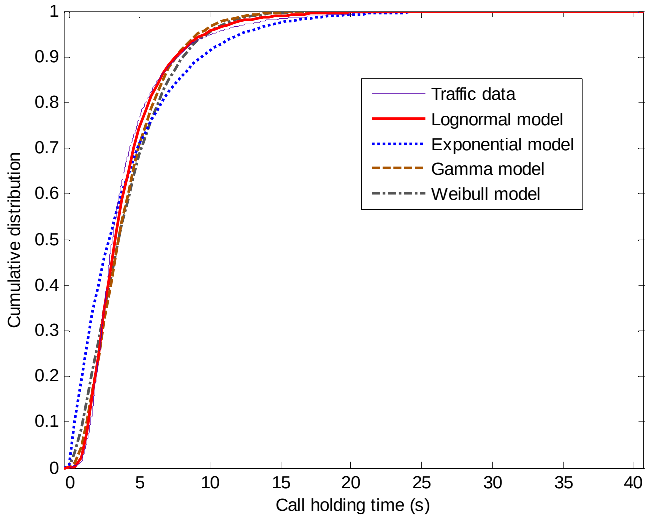

Fig. 8. Cumulative distribution function of the call holding times and comparison with exponential, Weibull, lognormal, and gamma distributions.

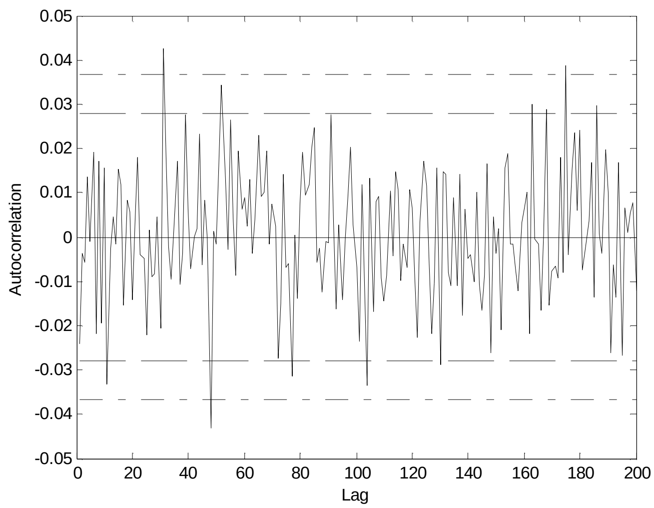

Fig. 9. Call holding time autocorrelation plot (up to lag 200) with 95% and 99% confidence intervals.

Table 7 shows the estimates of *H* for all hourly traces of call holding times, including the unreliable one (indicated by *). All estimates of H are close to 0.5, which implies that call holding times are not LRD and are uncorrelated. This is in agreement with the autocorrelation plot shown in Fig. 9.

4.4 Discussion and Comparison of Results from the Three Years

Erlang B and C models, which are used to model traffic in circuit-switched networks, are based on independent exponentially distributed call inter-arrival and holding times [10]. Our analysis of the traffic from the E-Comm network indicated that neither call holding nor call inter-arrival times fit the exponential distribution. In the case of call inter-arrival times, Weibull and gamma distributions are more suitable models. In [2], only the exponential distribution was identified as the best fit. Weibull and gamma distributions

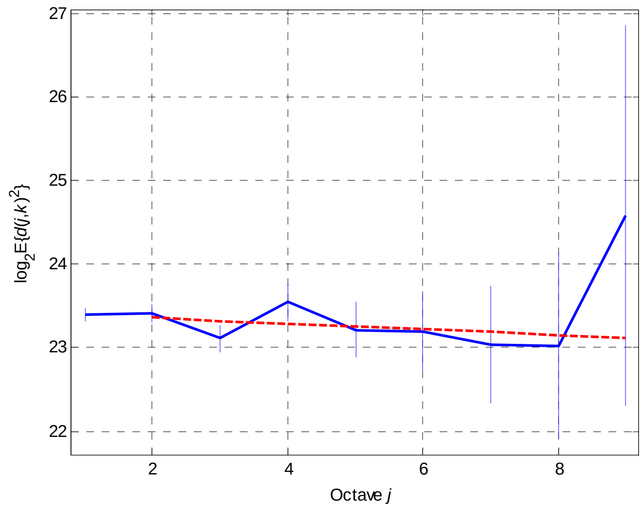

Fig. 10. Logscale diagram of the call holding times. Dashed line is the linear regression line with slope α. Vertical lines represent the 95 % confidence intervals around the estimates of $\log_2 E\{d(j,k)^2\}$.

Table 7. Estimates of *H* for the hourly traces of call holding times.

2001		2002		2003	
Day/hour	*H*	Day/hour	*H*	Day/hour	*H*
02.11.2001 15:00–16:00	0.493	01.03.2002 04:00–05:00	0.490	26.03.2003 22:00–23:00	0.483
01.11.2001 00:00–01:00	0.471	01.03.2002 22:00–23:00	0.460	25.03.2003 23:00–24:00	0.483
02.11.2001 16:00–17:00	0.462	01.03.2002 23:00–24:00	0.489	26.03.2003 23:00–24:00	0.463 *
01.11.2001 19:00–20:00	0.467	01.03.2002 00:00–01:00	0.508	29.03.2003 02:00–03:00	0.526
02.11.2001 20:00–21:00	0.479	02.03.2002 00:00–01:00	0.503	29.03.2003 01:00–02:00	0.466

are more general representations of the exponential distribution and their greater versatility makes them a better fit for the call inter-arrival times. Nevertheless, all three distributions (exponential, Weibull, and gamma) follow closely the distribution of the traffic data, as indicated in Figs. 3 and 4. Lognormal distribution best fits the call holding times, which agrees with the findings from [2]. As indicated in Figs. 7 and 8, the distribution of call holding times significantly deviates from the exponential distribution. The parameters of the best fitting distributions for the call inter-arrival and call holding times from 2003 are shown in Table 8. The analytical expressions for the distributions are given in Table 9, where $I(x)$ is an incomplete beta function and $\Gamma(a)$ is the gamma function.

Table 10 shows the average call inter-arrival and holding times for each of the busy hour traces. The average call inter-arrival times are 1.04 s, 0.86 s, and 0.84 s for the 2001, 2002, and 2003 dataset, respectively. The corresponding average call holding times are 3.91 (2001),

Table 8. Parameters of the best fitting distributions for the call inter-arrival and call holding times for the busy hours.

Busy hour	Distribution						
	Call inter-arrival times					Call holding times	
	Exponential	Weibull		Gamma		Lognormal	
	μ	a	b	a	b	μ	σ
02.11.2001, 15:00–16:00	0.9714	0.9785	1.0175	1.0326	0.9407	1.0913	0.6910
01.11.2001, 00:00–01:00	0.9711	0.9907	1.0517	1.0818	0.8977	1.0801	07535
02.11.2001, 16:00–17:00	1.0337	1.0651	1.0826	1.1189	0.9238	1.1432	0.6803
01.11.2001, 19:00–20:00	1.0868	1.1195	1.0800	1.1295	0.9622	1.1261	0.7040
02.11.2001, 20:00–21:00	1.1170	1.1373	1.0459	1.0760	1.0381	1.0893	0.7050
01.03.2002, 04:00–05:00	0.8121	0.8313	1.0603	1.1096	0.7319	1.1746	0.6671
01.03.2002, 22:00–23:00	0.8354	0.8532	1.0542	1.0931	0.7643	1.1157	0.6565
01.03.2002, 23:00–24:00	0.8620	0.8877	1.0790	1.1308	0.7623	1.1096	0.6803
01.03.2002, 00:00–01:00	0.9082	0.9266	1.0509	1.0910	0.8325	1.1334	0.6585
02.03.2002, 00:00–01:00	0.9160	0.9400	1.0686	1.1122	0.8236	1.1610	0.6680
26.03.2003, 22:00–23:00	0.7336	0.7475	1.0475	1.0910	0.6724	1.1838	0.6553
25.03.2003, 23:00–24:00	0.8492	0.8622	1.0376	1.0762	0.7891	1.1737	0.6715
26.03.2003, 23:00–24:00	0.8546	0.8579	1.0092	1.0299	0.8292	1.1704	0.6696
29.03.2003, 02:00–03:00	0.8711	0.8918	1.0617	1.0973	0.7939	1.1929	0.6437
29.03.2003, 01:00–02:00	0.8821	0.8970	1.0425	1.0705	0.8240	1.2373	0.6214

3.96 (2002), and 4.13 (2003). It can be observed that the average call inter-arrival times decreased and call holding times increased when going from 2001 to 2002 to 2003. This implies an increase in the traffic volume and increased network utilization.

Estimates of H for the traces of both call inter-arrival and call holding times from 2001, 2002, and 2003 are shown in Fig. 11. The estimates for each year are sorted in ascending order. The horizontal axis represents the rank (position in the sorted series of estimates) and the vertical axis shows the values of the estimates. Estimates of H for the traces of call holding times from the three years are very close to each other. The difference between the largest and the smallest estimate is approximately 0.05. Estimates of H for the call inter-arrival times exhibit greater variability. The difference between the largest and the smallest estimate is approximately 0.2. Comparing the Hurst parameter estimates of the traces from the various datasets (2001, 2002, and 2003), there are no large differences or an increasing or decreasing trend across the years. Rather, our findings indicate that the Hurst parameter may be regarded as an invariant characteristic of the busy hour traffic from the E-Comm network for the datasets from 2001, 2002, and 2003.

Table 9. The best fitting distributions.

Expression	Name: parameters
$f(x)=e^{-x/\mu}/\mu$	exponential: μ
$f(x)=ba^{-b}x^{b-1}e^{-(x/a)^b}I_{(0,\infty)}(x)$	Weibull: a, b
$f(x)=x^{a-1}e^{-(x/b)}/(b^a\Gamma(a))$	gamma: a, b
$f(x)=e^{-(\ln x-\mu)^2/(2\sigma^2)}/(x\sigma\sqrt{2\pi})$	lognormal: μ, σ

Table 10. Average call inter-arrival and call holding times from the hourly traces.

	2001		2002		2003	
	Day/hour	Avg. (s)	Day/hour	Avg. (s)	Day/hour	Avg. (s)
inter-arrival	02.11.2001 15:00–16:00	0.97	01.03.2002 04:00–05:00	0.81	26.03.2003 22:00–23:00	0.73
holding		3.78		4.07		4.08
inter-arrival	01.11.2001 00:00–01:00	0.97	01.03.2002 22:00–23:00	0.83	25.03.2003 23:00–24:00	0.85
holding		3.95		3.84		4.12
inter-arrival	02.11.2001 16:00–17:00	1.03	01.03.2002 23:00–24:00	0.86	26.03.2003 23:00–24:00	0.85
holding		3.99		3.88		4.04
inter-arrival	01.11.2001 19:00–20:00	1.09	01.03.2002 00:00–01:00	0.91	29.03.2003 02:00–03:00	0.87
holding		3.97		3.95		4.14
inter-arrival	02.11.2001 20:00–21:00	1.12	02.03.2002 00:00–01:00	0.91	29.03.2003 01:00–02:00	0.88
holding		3.84		4.06		4.25

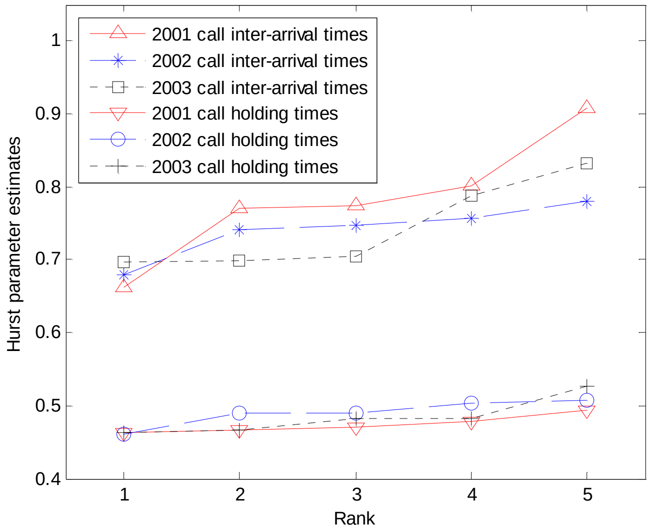

Fig. 11. Hurst parameter estimates of the busy hour traffic traces from 2001, 2002, and 2003.

5. CONCLUSIONS

In this paper, we analyzed busy hour voice traffic in a trunked, public safety wireless network from three consecutive years. We considered traffic from the cell with the largest capacity and that handles the majority of the calls. Our findings indicated that both Weibull and gamma distributions are suitable for modeling call inter-arrival times. Call holding times fit the lognormal distribution. Another important result of our analysis was that call inter-arrival times are long-range dependent. Therefore, traffic models that assume independent exponentially distributed call inter-arrival and holding times, such as Erlang B and C, may not produce reliable results if applied to trunked radio networks.

ACKNOWLEDGMENT

The authors thank D. Sharp from Planetworks and the management and technical staff at E-Comm for providing access to the activity data and technical support for data analysis. The authors also thank J. Song for helping with the data extraction and formatting.

REFERENCES

[1] N. Cackov, B. Vujičić, S. Vujičić, and Lj. Trajković, "Using network activity data to model the utilization of a trunked radio system," in *Proc. SPECTS*, San Jose, CA, July 2004, pp. 517–524.

[2] D. Sharp, N. Cackov, N. Lasković, Q. Shao, and Lj. Trajković, "Analysis of public safety traffic on trunked land mobile radio systems," *IEEE J. Select. Areas Commun.*, vol. 22, no. 7, pp. 1197–1205, Sept. 2004.

[3] J. Song and Lj. Trajković, "Modeling and performance analysis of public safety wireless networks," in *Proc. IEEE IPCCC*, Phoenix, AZ, Apr. 2005, pp. 567–572.

[4] E-Comm, Emergency Communication for SW British Columbia Incorporated [Online]. Available: http://www.ecomm.bc.ca.

[5] H. H. Hoang, R. Malhame, and C. Rosenberg, "Communication load and delay in multichannel land mobile systems for dispatch traffic: a queuing analysis," in *Proc. IEEE VTC*, Denver, CO, May 1992, pp. 773–777.

[6] P. Cohen, D. Haccoun, and H. H. Hoang, "Traffic analysis for different classes of land mobile communication systems," in *Proc. IEEE VTC*, Toronto, ON, May 1983, pp. 283-285.

[7] C. Jedrzycki and V. C. M. Leung, "Probability distribution of channel holding time in cellular telephony systems," in *Proc. IEEE VTC*, Atlanta, GA, Apr. 1996, pp. 247–251.

[8] F. Barcelo and J. Jordan, "Channel holding time distribution in public telephony systems (PAMR and PCS)," *IEEE Trans. Vehicular Technology*, vol. 49, no. 5, pp. 1615–1625, Sept. 2000.

[9] H. Chen and Lj. Trajković, "Trunked radio systems: traffic prediction based on user clusters," in *Proc. ISWCS*, Mauritius, Sept. 2004, pp. 76–80.

[10] G. Hess, *Land-mobile radio system engineering*. Norwood, MA: Artech House, 1993, pp. 249–286.

[11] M/A-Com EDACS [Online]. Available: http://www.macom-wireless.com/products/pve/edacs.asp.

[12] W. Leland, M. Taqqu, W. Willinger, and D. Wilson, "On the self-similar nature of Ethernet traffic (extended version)," *IEEE/ACM Trans. Networking*, vol. 2, pp. 1–15, Feb. 1994.

[13] K. Park and W. Willinger, "Self-similar network traffic: an overview," in *Self-similar Network Traffic and Performance Evaluation*, K. Park and W. Willinger, Eds. New York: Wiley, 2000, pp. 1–38.

[14] P. Abry, P. Flandrin, M. S. Taqqu, and D. Veitch, "Wavelets for the analysis, estimation, and synthesis of scaling data," in *Self-similar Network Traffic and Performance Evaluation*, K. Park and W. Willinger, Eds. New York: Wiley, 2000, pp. 39–88.

[15] D. Veitch, Matlab code for estimation of scaling exponents [Online]. Available: http://www.cubinlab.ee.mu.oz.au/~darryl/secondorder_code.html.

[16] D. Veitch and P. Abry, "A statistical test for time constancy of scaling exponents," in *IEEE Trans. Signal Processing*, vol. 49, no. 10, pp. 2325–2334, Oct. 2001.

[17] R. B. D'Agostino and M. A. Stephens, Eds., *Goodness-of-Fit Techniques*. New York: Marcel Dekker, 1986. pp. 63–93, pp. 97–145, pp. 421–457

[18] Industry Canada, Channel loading guidelines [Online]. Available: http://strategis.ic.gc.ca/epic/internet/insmt-gst.nsf/en/sf08015e.html.