A Novel Approach to the Estimation of the Long-Range Dependence Parameter

Houssain Kettani and John A. Gubner

***Abstract*—A new method to estimate the Hurst parameter of certain classes of random processes is presented. This method applies to Gaussian processes that are either exactly second-order self-similar or fractional ARIMA. The case of the former is of special interest because local area network traffic is well-known to be of this form. Confidence intervals and bias are obtained for the estimates using the new method. The new method is then applied to pseudo-random data and to real traffic data. The performance of the new method is compared to that of the widely-used wavelet method, which demonstrates that the former is much faster and produces much smaller confidence intervals of the long-range dependence parameter.**

***Index Terms*—Estimation, long-range dependence, network traffic, self-similarity.**

I. Introduction

IT IS NOW generally accepted that network traffic exhibits the features of long-range dependence and self-similarity [5], [10]–[15]. The parameter that measures these features is known as the Hurst parameter H, and many methods for estimating H have been proposed. For example, the following methods are described in the text by Beran [2]: rescaled adjusted range statistics (R/S) method, variance-time analysis, Higuchi's method, correlogram method, periodogram method, and Whittle estimator. More recent methods are the residuals of regression method due to Peng *et al.* [16], the wavelet method due to Abry and Veitch [1], and a method given by Chang and Chang [3]. See [17] for an empirical study of the estimates obtained using these methods. The analysis in [17] shows that the Whittle method is better than the classical methods. In [1], it is shown that the wavelet method is better than the Whittle method.

Among the above estimation methods, the only one that is not based on graphical analysis is the Whittle method. The only one that is based on the exact form of the covariance function is the variance-time method. And only the Whittle method and the wavelet method yield confidence intervals of the estimate. By far, the wavelet method is the most widely used. Thus, performance analysis of the proposed method in this brief will be compared to that of the wavelet method.

Manuscript received February 2005; revised July 2005. This paper was presented in part at the IEEE Conference on Local Computer Networks (LCN'02),Tampa, FL, November 6–8, 2002. This paper was recommended by Associate Editor G. Chen.

H. Kettani is with the Department of Computer Science, Jackson State University, Jackson, MS 39217 USA (e-mail: houssain.kettani@jsums.edu).

J. A. Gubner is with the Department of Electrical and Computer Engineering, University of Wisconsin, Madison, WI 53706 USA (e-mail: gubner@engr.wisc.edu).

Digital Object Identifier 10.1109/TCSII.2006.873828

As LAN traffic is generally accepted as *exactly* second-order self-similar [10], which specifies the form of the covariance function, we propose a new method for estimating H that exploits this structure. The new method is much faster and yields smaller confidence intervals than the wavelet method.

The remainder of the brief is organized as follows. Section II presents mathematical definitions and properties that will be used throughout the brief. In Section III, we present the proposed method for estimating the Hurst parameter of exactly second-order self-similar processes. We then apply this method to artificial and real traffic data in Section IV to check its performance, and compare it with that of the wavelet method. We present concluding remarks and further research directions in Section VI. In the Appendix we modify the ideas introduced in the body of the brief to apply them to fractional ARIMA processes.

II. Preliminaries

Let X_i denote the number of bits, bytes, or packets seen during the ith interval. We say that X_i is *second-order stationary* if its mean $E(X_i)$ does not depend on i and if the autocovariance function

$$E[(X_i - E(X_i))(X_j - E(X_j))]$$

depends on i and j only through their difference $k = i - j$, in which case we write

$$\gamma(k) = E\left[\left(X_{i+k} - E(X_{i+k})\right)\left(X_i - E(X_i)\right)\right].$$

The variance of the process is $\sigma^2 = \gamma(0) = E[(X_i - E(X_i))^2]$, and the autocorrelation function is $\rho(k) = \gamma(k)/\sigma^2$.

A second-order stationary process is said to be *exactly second-order self-similar* with Hurst parameter $H \in (0,1)$ if

$$\gamma(k) = (\sigma^2/2)(|k+1|^{2H} - 2|k|^{2H} + |k-1|^{2H})$$

or equivalently

$$\rho(k) = \frac{1}{2}\left(|k+1|^{2H} - 2|k|^{2H} + |k-1|^{2H}\right). \quad (1)$$

If X_i is a Gaussian process, it is known as *fractional Gaussian noise*.

III. Exactly Second-Order Self-Similar Processes

In this section, we develop a new estimator of the Hurst parameter of an exactly second-order self-similar process. Since X_i is exactly second-order self-similar, we have from (1) that

$$\rho(1) = 2^{2H-1} - 1.$$

We can solve for H to get

$$H = \frac{1}{2}\left[1 + \log_2\left(1 + \rho(1)\right)\right]. \quad (2)$$

Given observed data $X_1, \ldots, X_n$, let

$$\hat{\mu}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$$
$$\hat{\gamma}_n(k) = \frac{1}{n}\sum_{i=1}^{n-k}(X_i - \hat{\mu}_n)(X_{i+k} - \hat{\mu}_n)$$
$$\hat{\sigma}_n^2 = \hat{\gamma}_n(0)$$

and

$$\hat{\rho}_n(k) = \frac{\hat{\gamma}_n(k)}{\hat{\sigma}_n^2} \quad (3)$$

denote the *sample mean*, the *sample covariance*, the *sample variance*, and the *sample autocorrelation*, respectively. Based on (2), we propose

$$\hat{H}_n = \frac{1}{2}\left[1 + \log_2\left(1 + \hat{\rho}_n(1)\right)\right] \quad (4)$$

as an estimate of the Hurst parameter. To assess the performance of the proposed estimate, we appeal to the following special case of Hosking's result [7, Th. 6 and 7].

Theorem 1: Let X_i be an exactly second-order self-similar Gaussian process, i.e., fractional Gaussian noise. Then, for large sample size n, $\hat{\rho}_n(1)$ has mean

$$\mu_n = \rho(1) - \left(1 - \rho(1)\right) n^{2H-2}$$

and

1) If $H \in (0, (3/4))$, then $\hat{\rho}_n(1)$ is approximately $N(\mu_n, \sigma_n^2)$ with

$$\sigma_n^2 = \frac{1}{n}\Bigg\{ \left(1 + 3\rho^2(1)\right) + 2\sum_{k=1}^{\infty}\Big[\left(1 + 2\rho^2(1)\right)\rho^2(k) + \rho(k-1)\rho(k+1) - 4\rho(1)\rho(k-1)\rho(k)\Big]\Bigg\}. \quad (5)$$

2) If $H = (3/4)$, then $\hat{\rho}_n(1)$ is approximately $N(\mu_n, \sigma_n^2)$ with

$$\sigma_n^2 = \left[2H(2H-1)\left(1-\rho(1)\right)\right]^2 \frac{\log n}{n}. \quad (6)$$

3) If $H \in ((3/4), 1)$, then the limiting distribution of $\hat{\rho}_n(1)$ has mean μ_n and variance given by

$$\sigma_n^2 = 2\left[H(2H-1)\left(1-\rho(1)\right)\right]^2 K_2(H) n^{4H-4} \quad (7)$$

where $K_2(H)$ is related to the variance of the modified Rosenblatt distribution and is given by

$$K_2(H) = \int_0^1\int_0^1 g^2(x,y)\,dx\,dy$$

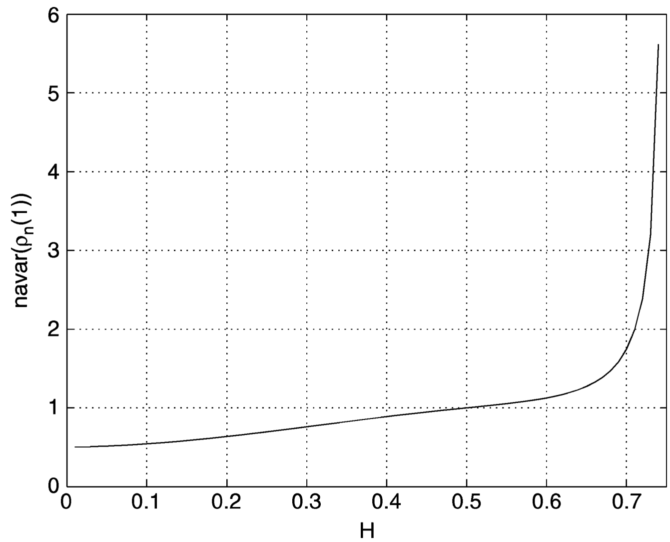

Fig. 1. Plot of $n\sigma_n^2$ in (5) as a function of the Hurst parameter H, in the second-order self-similar case.

where

$$g(x,y) = |x-y|^{2H-2} + \frac{1}{H(2H-1)} - \frac{1}{2H-1}\Big(x^{2H-1} + y^{2H-1} + (1-x)^{2H-1} + (1-y)^{2H-1}\Big).$$

A plot of $n\sigma_n^2$ in (5) (which does not depend on n) as a function of the Hurst parameter H, summing over $k = 1$ to 10^7 (instead of $k = 1$ to ∞ as in (5)) is given in Fig. 1. Now that we know $\hat{\rho}_n(1)$ is $N(\mu_n, \sigma_n^2)$

$$P\left(\left|\frac{\hat{\rho}_n(1) - \mu_n}{\sigma_n}\right| \le 1.96\right) = 0.95$$

i.e.,

$$\mu_n - 1.96\sigma_n \le \hat{\rho}_n(1) \le \mu_n + 1.96\sigma_n$$

holds with 95% probability. Using (4)

$$h_- \le \hat{H}_n \le h_+$$

where

$$h_\pm = \frac{1}{2}\left\{1 + \log_2\left[1 + \rho(1) - \left(1-\rho(1)\right)n^{2H-2} \pm 1.96\sigma_n\right]\right\} \quad (8)$$

also holds with 95% probability.

A. Comments on Confidence Intervals

For known H, the 95% confidence interval of the estimate $\hat{H}_n$ is $[h_-, h_+]$, with h_- and h_+ as in (8), with σ_n as in Fig. 1 if $H \in (0, 3/4)$, as in (6) if $H = 3/4$ and as in (7) if $H \in (3/4, 1)$. Let w_n denote the width of such intervals, i.e.,

$$w_n = h_+ - h_-.$$

A log–log plot of w_n versus the number of samples n is given in Fig. 2 for different values of H. It is remarkable to see the plot

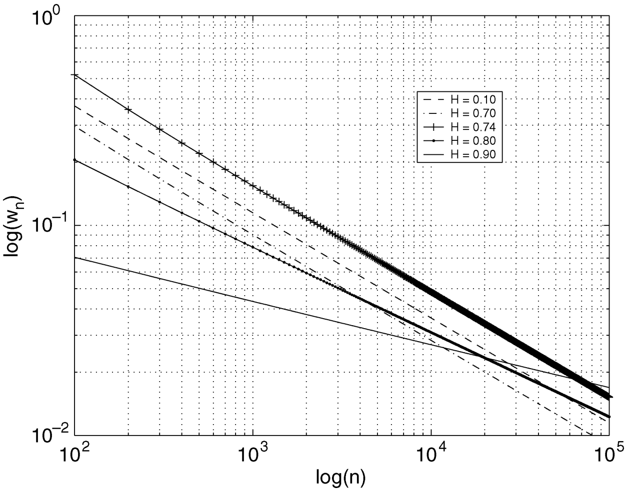

Fig. 2. Plot of the width of the 95% confidence intervals for different H values.

TABLE I
VALUES OF CONSTANTS a AND b IN (9) FOR EACH VALUE OF H

H	a	b
0.10	3.65	0.50
0.20	3.44	0.50
0.30	3.28	0.50
0.40	3.08	0.50
0.50	2.85	0.50
0.60	2.65	0.50
0.70	2.92	0.50
0.74	5.00	0.50
0.75	1.63	0.45
0.80	1.28	0.40
0.90	0.18	0.21

resembling a straight line for each value of H. Thus, the width w_n can be written as

$$w_n \approx a n^{-b} \tag{9}$$

where a and b are constants for fixed H. The values of these constants are given in Table I. It is interesting to note that the width w_n is upper bounded by the w_n at $H = 0.74$. Hence, in the case when H is not known (which is the typical case with real data), we choose the confidence interval centered around $\hat{H}_n$ with width

$$w_n = \frac{5}{\sqrt{n}}. \tag{10}$$

B. Summary of the Algorithm

In what follows, we present a summary of the new method.

- Let $X_1, X_2, \ldots, X_n$ be a realization of a Gaussian second-order self-similar process.
- Compute $\hat{\rho}_n(1)$ as in (3).
- Compute $\hat{H}_n$ as in (4), which is the estimated Hurst parameter.
- The 95% confidence interval of H is centered around the estimate $\hat{H}_n$ with width as in (10).

IV. Illustrative Examples

For each value of $H = 0.10, 0.20, \ldots, 0.90$, we generate 100 realizations of a fractional Gaussian noise. The length of each realization is $n = 4000$ points. For a given estimation method, we obtain 100 estimated values of H. Call these estimates $\hat{H}_n^{(k)}$, $k = 1, 2, \ldots, 100$. We compute their sample mean. We also provide the theoretical and empirical 95% confidence intervals of the estimates $\hat{H}_n$ and $\hat{H}_n^w$ from the proposed method and the wavelet method, respectively. The result of the application of the new method and the wavelet method to these data sets is given in Tables II and III, respectively.

TABLE II
RESULTS OF EMPIRICAL AND THEORETICAL STUDY OF NEW METHOD USING 100 INDEPENDENT REALIZATIONS

H	Mean of $\hat{H}_n$	Theoretical CI_o	Empirical CI_o
0.10	0.10	[0.06,0.14]	[0.07,0.13]
0.20	0.20	[0.17,0.23]	[0.17,0.23]
0.30	0.30	[0.27,0.33]	[0.27,0.32]
0.40	0.40	[0.37,0.43]	[0.37,0.43]
0.50	0.50	[0.48,0.52]	[0.47,0.52]
0.60	0.60	[0.58,0.62]	[0.58,0.63]
0.70	0.70	[0.67,0.73]	[0.68,0.72]
0.75	0.74	[0.72,0.77]	[0.72,0.76]
0.80	0.79	[0.77,0.81]	[0.77,0.82]
0.90	0.87	[0.86,0.90]	[0.85,0.90]

TABLE III
RESULTS OF EMPIRICAL AND THEORETICAL STUDY OF THE WAVELET METHOD USING 100 INDEPENDENT REALIZATIONS

H	Mean of $\hat{H}_n^w$	Theoretical CI_w	Empirical CI_w
0.10	0.00	[-0.10,-0.02]	[-0.05,0.06]
0.20	0.16	[0.07,0.15]	[0.10,0.22]
0.30	0.28	[0.21,0.28]	[0.21,0.34]
0.40	0.39	[0.33,0.40]	[0.32,0.44]
0.50	0.50	[0.44,0.52]	[0.44,0.56]
0.60	0.60	[0.55,0.62]	[0.55,0.67]
0.70	0.70	[0.65,0.73]	[0.65,0.75]
0.75	0.76	[0.73,0.85]	[0.70,0.82]
0.80	0.81	[0.75,0.83]	[0.75,0.87]
0.90	0.91	[0.86,0.93]	[0.84,0.95]

From both Tables II and III, it is observed that the confidence intervals obtained through the new method CI_o are narrower than those obtained through the wavelet method. The width of the empirical confidence intervals for the optimization method is about 0.05 versus 0.11 to 0.13 for those obtained through the wavelet method.

The theoretical and empirical confidence intervals are almost the same for the new method. This similarity holds even for $H \geq 0.80$, where we assumed normality although we knew that the distribution of $\hat{\rho}_n(1)$ is not normal. On the other hand, for the wavelet method, the empirical confidence intervals are considerably wider than theoretical ones, with the difference in width getting as high as 0.11 versus 0.07 for $H = 0.90$.

For $H < 0.80$, we see that the mean of the estimated Hurst parameter obtained by the new method $\hat{H}_o$ is the same as the true value H. For $H < 0.40$, the mean of the estimated Hurst parameter obtained by the wavelet method $\hat{H}_w$ is far from the true H and the latter does not fall in the 95% confidence interval. For $0.40 \leq H \leq 0.70$, the mean of $\hat{H}_w$ is closer to H and the theoretical confidence intervals contain the true value.

For $H = 0.80$, the new method under estimates the true value, with the mean of the estimates $\hat{H}_n$ is 0.79. The wavelet method, on the other hand, over estimates the true Hurst parameter value by the same quantity, namely the mean of the estimates $\hat{H}_n^w$ is 0.81. For $H = 0.90$, the estimates produced by the wavelet method over estimate the true value by the same quantity, namely the mean of the estimates $\hat{H}_n^w$ is 0.91. The new method, on the other hand, under estimates H, with the mean of the estimates $\hat{H}_n$ is 0.87. In this case, on average, $\hat{H}_n^w$ are closer to the true value than $\hat{H}_n$. However, the empirical confidence intervals of the wavelet method are much larger than those of the new method, with the width of the former is almost double the latter. It is also worth noting that for $H > 0.20$, the confidence intervals of the estimates obtained by the new method are contained in those of the wavelet method.

The number of *flops* is 2.5×10^4 for the new method and 1.3×10^7 for the wavelet method. Thus, the former is 520 times faster than the latter. In general, it is apparent that the new method gives more accurate and reliable results and is much faster than the wavelet method.

We next consider real data to test both methods. For this purpose, we investigate Ethernet measurements for a local area network traffic at Bellcore, Morristown, NJ [10]. From this data we extract a subset with length $n = 8475$ representing the amount of traffic observed each 100 ms. Passing the Bellcore data through the wavelet method gives $\hat{H}_n^w = 0.79$ with 95% confidence interval $CI_w = [0.75, 0.82]$. The new method, on the other hand, results in the value $\hat{H}_n = 0.81$ with confidence interval $CI_o = [0.78, 0.84]$. The variance-time, R/S, and periodogram methods resulted in $\hat{H} = 0.80, 0.79, 0.82$, respectively [10].

In short, it is clear that the new method and the wavelet method give close estimates with both real and artificial data. It is also noted that in all cases considered in this section, the new method's estimates fall in the 95% confidence intervals of the wavelet method. Moreover, the confidence intervals of the new method are contained in those of the wavelet method. Finally, the new method was shown to be much faster than the wavelet method.

V. Comments on LRD Testing

The ideas introduced in this brief can be generalized to provide a new method for testing whether a process is long-range dependent or not. This method uses the structure of the autocorrelation function of the model that we would like to fit the data to. We estimate the autocorrelation function of the data as described in (3), and then apply a curve-fitting criterion. To this end, the following *error function* is defined as

$$E_K(\beta) \doteq \frac{1}{4K}\sum_{k=1}^{K}\{\rho(k)-\hat{\rho}_n(k)\}^2 \tag{11}$$

where $\rho(k)$ denotes the autocorrelation function of the model with parameter β that we would like to fit the data to, $\hat{\rho}_n(k)$ is the sample autocorrelation function of the data, and K is the largest value of k for which $\hat{\rho}_n(k)$ is to be computed to reduce edge effects.

We expect $E_K(\beta)$ to be close to zero if X_i is close to the model. The estimated parameter $\hat{\beta}$ is chosen so that $E_K(\hat{\beta})$ is the minimum of the error function over the appropriate range of the parameter. It is easily seen that the highest $E_K(\hat{\beta})$ can be is 1. Thus, we consider that the prescribed model fits the process X_i if $E_K(\hat{\beta}) = e$, where e is "much smaller than 1".

This method can be used to estimate the Hurst parameter instead of stopping at the first lag of the sample autocorrelation function $\hat{\rho}_n(1)$ as introduced in this brief. However, while this would make the method more robust, it would also make it much more computationally intensive and we were unable to obtain the corresponding theoretical confidence intervals. See [8] for more detailed analysis on this method.

VI. Summary and Concluding Remarks

In this brief, we have presented a new tool to estimate the Hurst parameter in local area network traffic. The confidence intervals and bias of the estimates obtained by this new method are obtained. This new method is then applied to pseudo-random data and to real LAN traffic data. We compare the performance of the new method to that of the widely-used wavelet method. We demonstrated that the former is much faster and produces smaller confidence intervals of the Hurst parameter estimates. Furthermore, the confidence intervals of the estimates from the new method were shown to be contained in the confidence intervals of the wavelet method. Moreover, the new method was found to be much faster than the wavelet method.

Appendix
Fractional $\mathrm{ARIMA}(p, d, q)$ Processes

Two well-known processes that produce long-range dependence are fractional Gaussian noise processes and fractional ARIMA processes. The same ideas we introduced to estimate the long-range dependence parameter of fractional Gaussian noise can be applied to fractional $\mathrm{ARIMA}(p, d, q)$ processes. Thus, we devote this appendix to briefly illustrate such application.

The fractional $\mathrm{ARIMA}(p, d, q)$ process proposed in [4], [6] is an extension of the $\mathrm{ARIMA}(p, d, q)$ in the sense that d is allowed to take any real value in the interval $(-1/2, 1/2)$. Any fractional $\mathrm{ARIMA}(p, d, q)$ process can be expressed in terms of the *standard* fractional $\mathrm{ARIMA}(0, d, 0)$. The autocorrelation function of the latter is given by

$$\rho(k) = \frac{\Gamma(1-d)\Gamma(k+d)}{\Gamma(d)\Gamma(1+k-d)} = \prod_{i=1}^{k}\frac{k-i+d}{k-i+1-d}.$$

Now, suppose that X_i is known to be fractional $\mathrm{ARIMA}(0, d, 0)$. Then we have $\rho(1) = d/(1-d)$. Thus, we put

$$\hat{d}_n = \frac{\hat{\rho}_n(1)}{1+\hat{\rho}_n(1)} \tag{12}$$

to denote the estimated difference parameter of the process X_i. To assess the performance of the proposed estimate a similar result to Theorem 1 is obtained. To this end, we appeal to the following special case of Hosking's result [7, Th. 6 and 7].

Theorem 2: Let X_i be a fractional $\mathrm{ARIMA}(0, d, 0)$. Then, for large sample size n, $\hat{\rho}_n(1)$ has mean

$$\mu_n = \rho(1) - \frac{(1-2d)}{d(1-d)(1+2d)}\frac{\Gamma(1-d)}{\Gamma(d)}n^{2d-1}$$

TABLE IV
VALUES OF CONSTANTS a AND b IN (17) FOR DIFFERENT d VALUES

d	a	b
-0.40	14.88	0.50
-0.30	11.53	0.50
-0.20	9.32	0.50
-0.10	7.33	0.50
0.00	5.57	0.50
0.10	4.05	0.50
0.20	2.79	0.50
0.24	2.35	0.50
0.25	1.92	0.45
0.30	1.51	0.41
0.40	0.21	0.21

and

1) If $d \in (-1/2, 1/4)$, then $\hat{\rho}_n(1)$ is approximately $N(\mu_n, \sigma_n^2)$ with

$$\sigma_n^2 = \frac{2}{n}\left[\frac{1-2d}{1-d}\right]^2. \tag{13}$$

2) If $d = (1/4)$, then $\hat{\rho}_n(1)$ is approximately $N(\mu_n, \sigma_n^2)$ with

$$\sigma_n^2 = \left[\frac{2(1-2d)\Gamma(1-d)}{(1-d)\Gamma(d)}\right]^2 \frac{\log n}{n}. \tag{14}$$

3) If $d \in (1/4, 1/2)$, then the limiting distribution of $\hat{\rho}_n(1)$ has mean μ_n and variance given by

$$\sigma_n^2 = 2\left[\frac{(1-2d)\Gamma(1-d)}{(1-d)\Gamma(d)}\right]^2 K_2(d) n^{4d-2} \tag{15}$$

where $K_2(d)$ is related to the variance of the modified Rosenblatt distribution and is given by

$$K_2(d) = \int_0^1 \int_0^1 g^2(x,y)\,dx\,dy$$

where

$$g(x,y) = |x-y|^{2d-1} + \frac{1}{d(2d+1)} - \frac{1}{2d}\left(x^{2d} + y^{2d} + (1-x)^{2d} + (1-y)^{2d}\right).$$

With σ_n^2 given by Theorem 2, we have $d_- \le \hat{d}_n \le d_+$, with 95% probability, where

$$d_\pm = \frac{\mu_n \pm 1.96\sigma_n}{1 + \mu_n \pm 1.96\sigma_n}. \tag{16}$$

Now, for known d, the 95% confidence interval of the estimate $\hat{d}_n$ is $[d_-, d_+]$. Next, let w_n denote the width of such intervals, i.e., $w_n = d_+ - d_-$. A log–log plot of w_n versus the number of samples n for different values of d is similar to the one given in Fig. 2 and is therefore omitted. Thus, the width w_n can be written as

$$w_n \approx a n^{-b} \tag{17}$$

where a and b are constants for fixed d value. The values of these constants are given in Table IV.

It is interesting to note that the width w_n is upper bounded by the w_n at the value $d = -0.40$. Hence, in the case when d is not known, we choose the confidence interval centered around $\hat{d}_n$ with width

$$w_n = \frac{15}{\sqrt{n}}. \tag{18}$$

A. Summary of Algorithm

In what follows, we present a summary of the new method.

- Let $X_1, X_2, \ldots, X_n$ be a realization of a fractional $\mathrm{ARIMA}(0,d,0)$ process.
- Compute $\hat{\rho}_n(1)$ as in (3).
- Compute $\hat{d}_n$ as in (12), which is the estimated difference parameter,
- The 95% confidence interval of d is centered around the estimate $\hat{d}_n$ with width as in (18).

Finally, we note that simulation results comparing this algorithm with the wavelet method yielded similar conclusions to those drawn in Section IV.

ACKNOWLEDGMENT

The authors thank the reviewers for their constructive comments and valuable suggestions during the revision process.

REFERENCES

[1] P. Abry and D. Veitch, “Wavelet analysis of long-range dependent traffic,” *IEEE Trans. Inf. Theory*, vol. 44, no. 1, pp. 2–15, Jan. 1998.
[2] J. Beran, *Statistics for Long-Memory Processes*. New York: Chapman & Hall, 1994.
[3] Y. Chang and S. Chang, “A fast estimation algorithm on the hurst parameter of discrete-time fractional Brownian motion,” *IEEE Trans. Signal Process.*, vol. 50, no. 3, pp. 554–559, Mar. 2002.
[4] C. W. J. Granger and R. Joyeux, “An introduction to long-memory models and fractional differencing,” *J. Time-Ser. Anal.*, vol. 1, pp. 15–29, 1980.
[5] M. Grossglauser and J. Bolot, “On the relevance of long-range dependence in network traffic,” *Comput. Commun. Rev.*, vol. 26, no. 4, pp. 15–24, 1996.
[6] J. R. M. Hosking, “Fractional differencing,” *Biometrika*, vol. 68, pp. 165–176, 1981.
[7] ——, “Asymptotic distribution of the sample mean, autocovariances, and autocorrelations of long-memory time series,” *J. Enonometr.*, vol. 73, pp. 261–284, 1996.
[8] H. Kettani, “A Novel approach to the estimation of the long-range dependence parameter,” Ph.D. dissertation, University of Wisconsin, Madison, 2002.
[9] ——, “A novel approach to the estimation of the Hurst parameter in self-similar traffic,” in *Proc. 27th Ann. IEEE Conf. Local Comput. Netw.*, 2002.
[10] W. Leland, M. Taqqu, W. Willinger, and D. Wilson, “On the self-similar nature of ethernet traffic (extended version),” *IEEE/ACM Trans. Netw.*, vol. 2, no. 1, pp. 1–15, Jan. 1994.
[11] R. Morris and D. Lin, “Variance of aggregated web traffic,” in *Proc. IEEE INFOCOM '00*, Tel Aviv, Israel, 2000, pp. 360–366.
[12] K. Park, G. Kim, and M. Crovella, “On the relationship between file sizes, transport protocols, and self-similar network traffic,” in *Proc. 4th Int. Conf. Network Protocols*, Columbus, OH, 1996, pp. 171, 180.
[13] ——, “On the effect of traffic self-similarity on network performance,” in *Proc. SPIE Int. Conf. Performance Contr. Network Syst.*, Boston, MA, 1997, pp. 296, 310.
[14] K. Park and W. Willinger , *Self-Similar Network Traffic: An Overview Self-Similar Network Traffic and Performance Evaluation*, K. Park and W. Willinger, Eds. New York: Wiley, 2000.
[15] V. Paxson and S. Floyd, “Wide-area traffic: the failure of poisson modeling,” *IEEE/ACM Trans. Netw.*, vol. 3, no. 3, pp. 226–244, Mar. 1995.
[16] C. K. Peng, S. V. Buldyrev, M. Simons, H. E. Stanley, and A. L. Goldberger, “Mosaic organization of DNA nucleotides,” *Phys. Rev. E*, vol. 49, pp. 1685–1689, 1994.
[17] M. S. Taqqu, V. Teverovsky, and W. Willinger, “Estimators for long-range dependence: an empirical study,” *Fractals*, vol. 3, no. 4, pp. 785–788, 1995.